



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 130752

TO: Barba Kroma
Location: REM-2B692C18
Art Unit: 1638
Friday, September 03, 2004

Case Serial Number: 10/732721

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Thank you for using STIC services.

This Page Blank (uspto)

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .nrpm and .nrpn. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn.

Because they contain data that is confidential, the results of Pending database searches should not be left in the case.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 06:48:28 ; Search time 6301 Seconds
(without alignments)
11404.975 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcagcggtcacaatac.....ctcgttagtttggaacggcg 1658

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mu.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mu.*

41: em.htgo.other.*

*ed. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	68.8	4.1	2000	6	AX655393	AX655393 Sequence
c	64.6	3.9	150918	2	BX784041	BX784041 Danio rer
3	60.6	3.7	200337	9	AL590608	AL590608 Human DNA
4	59.8	3.6	78101	8	AF031569	AF031569 Zea mays
5	59.8	3.6	346296	8	AF090447	AF090447 Zea mays
6	56.8	3.4	78756	2	PFMAL6P1_13	Continuation (14 o
c	56.4	3.4	273275	3	AE014828	AE014828 Plasmodiu
c	56.2	3.4	1791	3	AF151389	AF151389 Dermatobi
c	55.8	3.4	1792	3	AF151388	AF151388 Dermatobi
c	55.8	3.4	328	11	BV074228	BV074228 S212P6542
11	55.8	3.4	125316	9	AC092635	AC092635 Homo sapi
12	55.8	3.4	335050	3	PFA929356	AL929356 Plasmodiu
13	55	3.3	205429	2	AC005506	AC005506 Plasmodiu
c	55	3.3	253132	3	AE014846	AE014846 Plasmodiu
c	55	3.3	313050	3	PFA929352	AL929352 Plasmodiu
16	54.6	3.3	83902	5	AL672161	AL672161 Zebrafish
c	54.4	3.3	200337	9	AL590608	AL590608 Human DNA
18	54.4	3.3	302156	3	AC116977	AC116977 Dictyoste
c	54.2	3.3	111882	3	AC115612	AC115612 Dictyoste
20	54	3.3	313050	3	PFA929352	AL929352 Plasmodiu
c	53.6	3.2	2750	6	A43608	A43608 Sequence 33
c	53.6	3.2	2750	6	AR116777	AR116777 Sequence
23	53.6	3.2	12779	3	AE001381	AE001381 Plasmodiu
24	53.6	3.2	110000	2	PFMAL7P1_04	Continuation (5 of
25	53.6	3.2	110000	2	PFMAL7P1_05	Continuation (6 of
26	53.6	3.2	114475	9	AC108082	AC108082 Homo sapi
c	53.6	3.2	196248	2	AC020725	AC020725 Homo sapi
28	53.6	3.2	251065	2	AC115145	AC115145 Rattus no
c	53.4	3.2	110000	2	PFMAL8P1_06	Continuation (7 of
c	53.4	3.2	250029	3	AE014820	AE014820 Plasmodiu
c	53.4	3.2	250029	3	AE014839	AE014839 Plasmodiu
c	53.2	3.2	7498	6	AX345159	AX345159 Sequence
33	53.2	3.2	176212	2	AC141778	AC141778 Apis mell
34	53.2	3.2	192929	2	AC005505	AC005505 Plasmodiu
35	53.2	3.2	250713	3	AE014850	AE014850 Plasmodiu
c	53	3.2	58697	2	AC144975	AC144975 Silurana
c	53	3.2	141752	2	BX537131	BX537131 Danio rer
c	53	3.2	234081	3	PFMAL4P2	AL035475 Plasmodiu
c	52.8	3.2	103343	9	AC108078	AC108078 Homo sapi
c	52.8	3.2	110000	2	PFMAL6P1_04	Continuation (5 of
41	52.8	3.2	111882	3	AC115612	AC115612 Dictyoste
c	52.8	3.2	114276	9	AC011355	AC011355 Homo sapi
c	52.8	3.2	250029	3	AE014816	AE014816 Plasmodiu
c	52.6	3.2	172497	2	BX276093	BX276093 Danio rer
c	52.6	3.2	193113	5	AL935197	AL935197 Zebrafish

ALIGNMENTS

RESULT 1

AX655393

LOCUS

DEFINITION

AX655393

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS

AX655393 Sequence 5263 from Patent WO03000898. linear PAT 22-MAR-2003

AX655393 GI:29158207

Oryza sativa

Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,

Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.

RESULT_2	
EX784041/c	
LOCUS	BX784041
DEFINITION	Danio rerio clone DKEYP-95D10, WORKING DRAFT SSEQUENCE, 12 unordered pieces.
ACCESSION	BX784041
VERSION	BX784041.1 GI:38304074
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio

```

misc_feature      fragment_chain:1"
36429..54463
/note="assembly_fragment:00422
fragment_chain:1"
54564..65123
/note="assembly_fragment:00239
fragment_chain:2"
65224..72411
/note="assembly_fragment:00131
fragment_chain:2"
72512..75095
/note="assembly_fragment:00059
fragment_chain:2"
75196..90315
/note="assembly_fragment:00319
fragment_chain:3"
90416..95884
/note="assembly_fragment:00076
fragment_chain:3"
95985..123970
/note="assembly_fragment:00925
fragment_chain:4"
124071..127449
/note="assembly_fragment:00103
fragment_chain:4"
127550..150918
/note="assembly_fragment:00560"

ORIGIN
Query Match      3.9%; Score 64.6; DB 2; Length 150918;
Best Local Similarity 51.2%; Pred. No. 0.0025;
Matches 151; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 13 CACAATACCAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTCTGTGAATCAAT 72
Db 127880 CAAAATCAATACCATTTTGTCTCAACAAATATTAGAGAGTGGGTGTAACGTGCATGATA 127821

QY 73 TCAGAGGATGAAGAAAAATCGCAAGAACCAATATTTTAAATGAATGATGCAATATAC 132
Db 127820 TTGTAATAGAAAGAAAAATGAAAAAGGTGATTTATTATTAATAAACCAATTAAC 127761

QY 133 AATTTAATTAACAATTAATGTAAGATTACATTTGTTAGTTTCATAGAAATCAATTTCTA 192
Db 127760 ATTATTAATAGCATATTTGCTATTATTAGATAATATATATATATATATATATATATA 127701

QY 193 GAGTCATAATAATGCGCTAAATTAATTTCTATTCTTTTCCTTAAGAAAAAAGACAGC 252
Db 127700 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 127641

QY 253 CCATTAAAGGACCATTAAGATGCGCGCTGCTCCATTTTATATATATAGATAT 307
Db 127640 ATATATAAATTAATTATAAATATATATATTTATTTTATTTATATATATATATATATAT 127586

RESULT 3
AL590608 Human DNA sequence from clone RP11-363F12 on chromosome 6, complete
DEFINITION sequence.
ACCESSION AL590608
VERSION AL590608.6 GI:14530781
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200337)
AUTHORS Almeida,J.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Jun 23, 2001 this sequence version replaced gi:14529950.

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>

RP11-363F12 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-363F12 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-363F12 is at 1 in this sequence.

The true left end of clone RP1-244F1 is at 198338 in this sequence.

FEATURES

Location/Qualifiers

1..200337

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="6"

/clone_lib="RPCI-11.2"

723..955

/note="MIR repeat: matches 7. .256 of consensus"

repeat_region

1255..1753

/note="MUT1D repeat: matches 5. .502 of consensus"

repeat_region

2302..2846

/note="MER77 repeat: matches 28. .618 of consensus"

repeat_region

4615..4996

/note="LIMA9 repeat: matches 5378. .5759 of consensus"

repeat_region

4996..6390

/note="LIPA7 repeat: matches 4697. .6123 of consensus"

repeat_region

6342..7074

/note="LIMA9 repeat: matches 5783. .6308 of consensus"

repeat_region

7186..7544

/note="THE1C repeat: matches 1. .371 of consensus"

repeat_region

9108..9141

/note="17 copies 2 mer aa 82% conserved"

repeat_region

9287..9936

/note="25 copies 2 mer tt 72% conserved"

repeat_region

10742..10832

/note="MIR repeat: matches 59. .144 of consensus"

repeat_region

13160..19307

/note="LIPA5 repeat: matches 11. .6145 of consensus"

repeat_region

22327..22384

/note="29 copies 2 mer tg 70% conserved"

repeat_region

22842..22971

/note="TIGGER1 repeat: matches 2151. .2279 of consensus"

repeat_region

22970..23399

/note="TIGGER1 repeat: matches 1967. .2392 of consensus"

repeat_region

23400..23532

/note="FLAM_C repeat: matches 1. .133 of consensus"

repeat_region

23533..25472

/note="TIGGER1 repeat: matches 1. .1967 of consensus"

repeat_region

26263..26589

/note="AluY repeat: matches 1. .311 of consensus"

```
repeat_region 29354..29515
/note="81 copies 2 mer at 64% conserved"
repeat_region 29441..29518
/note="3 copies 26 mer 80% conserved"
repeat_region 29738..30283
/note="21 copies 26 mer 61% conserved"
repeat_region 30044..30271
/note="114 copies 2 mer ta 63% conserved"
repeat_region 30239..30585
/note="Aluub repeat: matches 3..282 of consensus"
repeat_region 30727..37903
/note="L1Mx3 repeat: matches -1388..6131 of consensus"
repeat_region 38096..38328
/note="Alufo repeat: matches 60..296 of consensus"
repeat_region 39023..39070
/note="24 copies 2 mer tt 77% conserved"
misc_feature complement(40758..41162)
/note="match: STS: Em:HSPA31G7"
repeat_region 42772..43589
/note="L1Mx1 repeat: matches 5440..6313 of consensus"
repeat_region 44021..44382
/note="L1P repeat: matches 4662..4985 of consensus"
repeat_region 44383..44685
/note="AluY repeat: matches 1..303 of consensus"
repeat_region 44686..44934
/note="L1P repeat: matches 4985..5244 of consensus"
repeat_region 44956..45591
/note="L1P repeat: matches 5527..6162 of consensus"
repeat_region 45614..45918
/note="L1Mx8 repeat: matches 4952..6280 of consensus"
repeat_region 50859..50998
/note="MER46A repeat: matches 1..140 of consensus"
repeat_region 51192..51609
/note="L1Mx5 repeat: matches 5721..6166 of consensus"
repeat_region 51842..51877
/note="18 copies 2 mer ca 83% conserved"
repeat_region 53075..53224
/note="75 copies 2 mer ta 73% conserved"
repeat_region 53085..53240
/note="6 copies 26 mer 72% conserved"
repeat_region 53264..53308
/note="AluSp repeat: matches 150..194 of consensus"
repeat_region 53309..53458
/note="Aluub repeat: matches 86..236 of consensus"
repeat_region 58009..58307
/note="AluSc repeat: matches 1..307 of consensus"
repeat_region 61286..61410
/note="L2 repeat: matches 2608..2737 of consensus"
repeat_region 62527..62693
/note="MIR repeat: matches 97..262 of consensus"
repeat_region 62842..63369
/note="L1Mx2D repeat: matches 1..533 of consensus"
repeat_region 63375..63433
/note="L1Mx3 repeat: matches 1..306 of consensus"
repeat_region 68004..68307
/note="L1Mx3 repeat: matches 490..553 of consensus"
repeat_region 69202..70145
/note="L1Mx3 repeat: matches 337..1364 of consensus"
repeat_region 70206..70563
/note="L1Mx3 repeat: matches 1666..2032 of consensus"
repeat_region 72600..72749
/note="L1Mx1 repeat: matches 6162..6314 of consensus"
repeat_region 75551..75715
/note="L1Mx3 repeat: matches 2385..2208 of consensus"
repeat_region 75697..76004
/note="L1Mx4 repeat: matches 5006..5325 of consensus"
repeat_region 76091..76275
/note="L1Mx3 repeat: matches 5962..6152 of consensus"
repeat_region 76963..77076
/note="57 copies 2 mer tt 60% conserved"
repeat_region 77189..77705
/note="MER68A repeat: matches 6..559 of consensus"
repeat_region 78243..78536

/note="AluY repeat: matches 1..292 of consensus"
78889..79139
/note="L1EB1 repeat: matches 5894..6154 of consensus"
79671..80056
/note="THE1C repeat: matches 1..371 of consensus"
80188..80485
/note="L1R18B repeat: matches 220..603 of consensus"
80630..80891
/note="AluX repeat: matches 1..262 of consensus"
82491..82744
/note="MIR repeat: matches 2..261 of consensus"
85379..85469
/note="AluSc repeat: matches 202..292 of consensus"
85573..85875
/note="MER67C repeat: matches 410..709 of consensus"
85876..85923
/note="L1R29 repeat: matches 571..619 of consensus"
85931..86543
/note="MER41B repeat: matches 1..635 of consensus"
86581..86632
/note="MER39B repeat: matches 494..545 of consensus"
86752..86996
/note="AluY repeat: matches 7..312 of consensus"
87291..87416
/note="L1R29 repeat: matches 4..137 of consensus"
87449..87759
/note="MER67C repeat: matches 106..414 of consensus"
87781..87889
/note="BUR1 repeat: matches 7386..7485 of consensus"
87999..88707
/note="L1P16 repeat: matches 5443..6157 of consensus"
88710..88750
/note="L1R28 repeat: matches 13..53 of consensus"
88727..89132
/note="MER61C repeat: matches 14..431 of consensus"
89148..90930
/note="MER61-internal repeat: matches 2814..4549 of

Query Match 3.7%; Score 60.6; DB 9; Length 200337;
Best Local Similarity 50.9%; Pred.No.0.022;
Matches 144; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Qy 24 TCATCTCATGTCGATAAATTTTATAGTTAGATTACTTCTTGTAATCATCTTCAGAGGATGA 83
Db 29705 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 29764

Qy 84 AAAAAAATCGCAAGCAAAATATTTTAAUGAATGATGCAATATACAAATTTAAATTA 143
Db 29765 AATATAAATATATATATATATATATATATATATATATATATATATATATATATATAT 29824

Qy 144 CACAAATATGTAAGATTACATTTGTTAGTTTCATAGAAACAATTTCTAGAGTCATATA 203
Db 29825 TATAAATATATATATATATATATATATATATATATATATATATATATATATATATA 29884

Qy 204 ATGCCTAAATTAATAATCTATTCATTTTCCTTAAGAAAAAACGCCCATTAAGGGA 263
Db 29885 TTATATATATATATATATATATATATATATATATATATATATATATATATATATATA 29944

Qy 264 CCATTGAATGCGCGCTGCCCAATTTTATATATATATATATATATATATATATATATAT 306
Db 29945 ACATTATCTATAAATTTATATATATATATATATATATATATATATATATATATATA 29987

RESULT 4
AF031569 AF031569 78101 bp DNA linear PLN 18-MAR-2003
DEFINITION Zea mays 22-kDa alpha zein gene cluster, complete sequence.
ACCESSION AF031569
VERSION AF031569.1 GI:2832242
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 78101)

AUTHORS Song,R., Llaca,V., Linton,E. and Messing,J.

TITLE Sequence, regulation, and evolution of the maize 22-kD alpha zein

gene family

JOURNAL Genome Res. 11 (11), 1817-1825 (2001)

MEDLINE 21549079

PUBMED 11691845

REFERENCE 2 (bases 1 to 78101)

AUTHORS Llaca,V. and Messing,J.

TITLE Direct Submission

JOURNAL Submitted (27-OCT-1997) Waksman Institute, Rutgers, The State

University of New Jersey, P.O. Box 759, Piscataway, NJ 08855-0759,

USA

FEATURES

source

Location/Qualifiers

1..78101

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/note="cosmid contig; inbred line BSSS53"

1..39083

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="III.4H10"

/note="inbred line BSSS53"

25530..68272

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="III.5D8"

/note="inbred line BSSS53"

40872..73554

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="V.3B1"

/note="inbred line BSSS53"

41990..78101

/organism="Zea mays"

/mol_type="genomic DNA"

/db_xref="taxon:4577"

/chromosome="4"

/map="4s"

/clone="V.9C11"

/note="inbred line BSSS53"

1..541

/note="similar to Zea mays cl locus myb homolog CDS and

gpc4 promoter region"

1471..2270

/gene="azs22-1"

/note="22-kDa alpha zein 1"

/pseudo

5582..6382

/gene="azs22-2"

/note="22-kDa alpha zein 2"

/pseudo

9689..9766

/gene="azs22-3"

/note="22-kDa alpha zein 3; truncated"

/pseudo

10702..10770

/note="microsatellite"

/rpt_type=tandem

repeat_region

repeat_region

/rpt_unit="AT"

15639..15694

/note="microsatellite"

/rpt_type=tandem

/rpt_unit="AT"

18119..18919

/gene="azs22-4"

18119..18919

/gene="azs22-4"

/codon_start=1

/product="22-kDa alpha zein 4"

/protein_id="AAC01573.1"

/db_xref="GI:2832243"

/translation="WATKILSLALLALFASATNAFIPOCSLAPSSIIOTFLPVPVTS

MGFHPAVQAYRLQQAATAASVILQPIISLOQQSLAHLTIOTIATCQQOQFIPLASHLA

MVNPAYLQQQLASNPALANVAWQVQQQLQQLFLPALSQLAWNPAYLQQOQLLS

SSPLAVANAPYVLLQQQLQQLVLPALTOLVANPAAYLQQQLLPFQLTMSNSAAYLQQR

QQLNPLAVANPLVAAFLQQQLLPYQFSLINPVLRSRQPIVGGAI"

23325..24125

/gene="azs22-5"

/note="22-kDa alpha zein 5"

/pseudo

27044..28675

/rpt_family="PREM-1 retroelement-related sequence"

/rpt_type=dispersed

30446..31840

/codon_start=1

/product="hypothetical protein"

/protein_id="AAC01574.1"

/db_xref="GI:2832244"

/translation="MVQIVVHRAKTSDFLVKKEAVSSTPQVNPVPVSGTAAVHAAR

QHTAOTSRTSPRAAGALLAARELLRHPSSSTASPGAMKQWDDVNRQLDMAHSTST

RSRQSSHROYEASTSMRSPSVRGQNTDLAVELNHRAGADAVLSLACERRONIE

GRNLDDFAEVVPPQAPMGTRSQAGVPLAGVCAALVDHLRAASWSPKRPPLPEKIDG

TSNSEFLQVYVITAAAGTSQAMATYFHVALSFGARTLNLNLSGPGSIYSWEELIAR

FVANFASAYQQHGVGVEHLHAVRQEPRETRMFISFTKVGQTIPTISDAISIATPRQG

VRDEKMLEKLATHDVEIVPTLLADKCAAAEGRAEHLAPQTGATGSGSGAI PRDG

KKKKKDRDHOKPRSTALVAVGTGGRGNRPRPQRGNSGSCVPHPNGRINTVECR

ELIDLAKVSRHE"

36867..37673

/gene="azs22-6"

/note="22-kDa alpha zein 6"

/pseudo

40977..42480

/gene="azs22-7"

/note="22-kDa alpha zein 7"

/pseudo

41499..42213

/note="insertion"

complement(48352..51063)

/note="similar to retrovirus-related POL polypeptide

sequence"

complement(49176..49850)

/codon_start=1

/product="hypothetical protein"

/protein_id="AAC01575.1"

/db_xref="GI:2832245"

/translation="MQYLATIRSLERQPKFTLQHVDRANKNEADALAKAARGAALP

SDVFYHVTGTPAFRSPGELQINDSEGHRIVNLIMTEWRAPITLFLQYIYPTDINE

AKCLHRSRDFALIEQLYKKGVSQMLKCVTEGQILREVHSGTCGSHAGPRALA

AKVIHQGFYWPAMICAAANVRVTRSCACQKFSRSGSPSQFTKLIHTWPLQRMGLDVI

GELPTA"

52998..53798

/gene="azs22-8"

52998..53798

/gene="azs22-8"

/codon_start=1

/product="22-kDa alpha zein 8"

/protein_id="AAC01576.1"

/db_xref="GI:2832246"

/translation="WATKILALLALLFVSATNAFIIPQCSLAPSAIIPQLRPPVTS

MGFHLAVQAYRLQQAASVILQPIINQIQSSLAHLTIQTATCQQOQFLPALSQLD

VVNVPAYLQQQLASNPALANVAAYQQOQLQQLFLPALSQLAWNPAYLQQOQLLS

SSPLVGNAPTYLQOQLQIIVPALTOAVANPAAYLQQLLPFNQLTVNSAGYLQOR
QQLNPLAVNELVTAFLQOQLLPYSQFSIMNPALSWQPIVGGALF"
58815..58610
/gene="azs22-9"
/note="22-kDa alpha zein 9"
/pseudo
complement(67340..68552)
/note="similar to Zea mays globulin-1 promoter"
misc_feature
68037..68563
/note="similar to Zea mays globulin-1 promoter"
misc_feature
71597..72372
/note="similar to Zea mays zrp2 promoter"
gene
75093..75893
/gene="azs22-10"
CDS
75093..75893
/gene="azs22-10"
/codon start=1
/product="22-kDa alpha zein 10"
/protein_id="AAC01577.1"
/db_xref="GI:2832247"
translation:"MATKILALLALLFVSATNFAFIPOCSLAPSAIIPOFLPPVTS
MGHEHSAVQNRQLQALAAVSLQPIAQLOOQSLAHITIOAITTQOQQQFLPALSHIA
MVNPAAYLQOQLLALANVAVNQOQLOOFLPALSLAMVNPAAVYVQOQQLIS
SSPLVGNAPTYLQOQLQIIVPALTOAVANPAAYLQQLLPFNQLTVNSAAYLQOR
QQLNPLVIANPLVAFLQOQLLPYNQFSIMNPVLARQQPIVGGALF"

ORIGIN

Query Match 3.6%; Score 59.8; DB 8; Length 78101;
Best Local Similarity 62.3%; Pred. No. 0.027;
Matches 160; Conservative 0; Mismatches 92; Indels 5; Gaps 4;
QY 942 AGTAGACCCAGTCGTGATAACGTGTTAAGAACCCCTGTACCGAGTGTAGTCCAAAGG 1001
Db 60320 AATGGAGCATTCGTGCTGATAACGTGTTAAGAACCCCTGTACTAGTGTAGTCCAAAGG 60379
QY 1002 CTCGCTTAACCT- GAATAATAGTAGACACAAAGGACAGAGTGTAGAGGAACTGATT 1059
Db 60380 CTTCTTAATAGAAATAGAAATCATAAACAGGACAGACACAAATGTATAGGAGAACTGTT 60439
QY 1060 CTTTGTACTATATGTCGTCTCTCAAGAGTTACATGATATGGGATCTCCTCTCTAT 1119
Db 60440 ATTCTTATTCATGTATT-CAGCTCANTGGTTACAAACAAATGGTGTCTCCTCTCTATAT 60498
QY 1120 TTATAGACAAAACCTAGGTTTTCAGGCATATGGGCCACATAGGCCTTCTCGGCCCAAGAAA 1179
Db 60499 ATAGGCT-AGAATTAGGTTTCGGCTAGATGGCCACATCGGCCTTCGGGGCCCAA-ACT 60556
QY 1180 GGTTCCTTAACACTACC 1196
Db 60557 GGTTCCTTAACACTCCC 60573

RESULT 5
AF090447
LOCUS
DEFINITION
Zea mays 22 kDa alpha zein gene cluster, complete sequence.
ACCESSION
AF090447 AF105716
VERSION
AF090447.2 GI:13606087
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 346296)
Song,R., Liaca,V., Linton,E. and Messing,J.
Sequence, regulation, and evolution of the maize 22-kD alpha zein
gene family
Genome Res. 11 (11), 1817-1825 (2001)
JOURNAL
MEDLINE
21549079
11691845
PUBMED
REFERENCE
2 (bases 80292 to 116863)
Song,R., Liaca,V. and Messing,J.

Direct Submission
Submitted (04-SEP-1998) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
3 (bases 1 to 65155)
Liaca,V., Lou,A., Young,S. and Messing,J.
Direct Submission
Submitted (10-NOV-1998) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
4 (bases 1 to 346296)
Song,R., Liaca,V. and Messing,J.
Direct Submission
Submitted (12-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Sequence update by submitter
5 (bases 1 to 346296)
Song,R., Liaca,V. and Messing,J.
Direct Submission
Submitted (24-APR-2001) Waksman Institute, Rutgers University, 190
Frelinghuysen Road, Piscataway, NJ 08854-8020, USA
Amino acid sequence updated by submitter
On or before Apr 12, 2001 this sequence version replaced
gi:4416300, gi:4140643.
FEATURES
source
1..346296
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid II.2E10"
/notes="inbred line BSS53"
1..65155
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid III.3C12; V.9D7"
/notes="inbred line BSS53"
1261..185036
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="BAC 204"
/notes="inbred line BSS53"
80292..116863
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="cosmid II.2E10"
/notes="inbred line BSS53"
172910..346296
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="4"
/map="4s"
/clone="BAC 171"
/notes="inbred line BSS53"
join(<1256..1397,2128..>2300)
/product="hypothetical protein"
join(1256..1397,2128..2300)
/codon start=1
/product="hypothetical protein"
/protein_id="AAD20309.1"
/db_xref="GI:4416304"
/translation="MWKLGKVLVLEHDDTDADRSSEKKLDVGFTEETTEQESTGVRYW
KAGSSATVPSSAQKVFMTVMRMEFFSLKCLPHSIKDNQNTGWTFRVNDGTELIRLQ
IR"

```

mRNA
join(<3179, .4286,5020, .5155,5266, .5443,5892, .>6020)
/product="hypothetical protein"
CDS
join(3179, .4286,5020, .5155,5266, .5443,5892, .6020)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20310.1"
/db_xref="GI:4416305"
/translation="MVQKGTVEGSDTSVLEQTQTVGASVSNLAQCCTCSAVFGDD
KEMANCKVEHSSAGRTVTVASGENTHTAGCGGCGSNCGPMMAEDPKAGNAKSGG
CGSGCGGCGGAGCTMANEGQSRSGVGTLLSNAMGEGQQAHSAGCGSGCGGG
GSMVTEGSKANNAMSGCGCGCGGCGGMLIEVSKANNAMSGCGCGCGGCGSGM
VNEGSKANNAMSGCGCGCGGCGGALFNASTAAGELKNNKSAAGCGCGGCGGGG
CGSNALIGFKNAHAKSGCGCGCGGCGGTLFSSAAAGQSRSGCGCGCGCGCGS
GSMVAEGNGHAKSGCGCGCGGCGGCGGATTLTSLMPGADVVAEESGTA
RCEELVGPAPTPPTAFITFYGGRKASLQRFACGDCGGGELSPFLMPTVETMVLHN
YIAPDSRLHAKDIYNPOHDFVETCCCSSTSPWLAASLQNAFEGTKBQTAATQ
TCCCCFQK"
join(<1112, .11154,11262, .11396,12254, .12344,14272, .14410,
14838, .14890,18185, .18230,18589, .>18657)
/product="hypothetical protein"
CDS
join(1112, .11154,11262, .11396,12254, .12344,14272, .14410,
14838, .14890,18185, .18230,18589, .18657)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20311.1"
/db_xref="GI:4416306"
/translation="MDVGMWMTATRVDERSRHAGDGLTQMSGVATGEAETHGGG
EEVEGVQDMAPQREVLGRDMIGRTSTCSVGTWSVAPPRGLARYITEPLLEMSF
DHRPCYAGLGGVEIPEKRTKVLDPSLPHYCESATNLYEALDNRDGLSRQLLRSI
GMEDPKVTLVKNQSAISLSIKDPQAAK"
15142, .17012
/note="similar to cell division protein FTSz"
22068, .22142
/rpt_type=dispersed
24303, .24307
/note="sequence target duplication"
24306, .31758
/rpt_family="zeon-1 retroelement"
24308, .24976
/rpt_type=dispersed
<25532, .>26569
/product="gag protein"
25532, .26569
/codon_start=1
/product="gag protein"
/protein_id="AAD20308.1"
/db_xref="GI:4416303"
/translation="MSYQATVSSYGNATVMAKSFVMVRSVAQTWYSSLRPGTITSW
QKLDMLTSGFGTQKPTAQLFCQDHEYLQAVVRFRLRLRAQPTVPNEIVI
EAMIKGLRPGSAQYFARKPOTLEKLQKMDXEYIRADNFRREERAFRSEMTRGF
GGFYPRHVRSHNDSQSQOQRPQCSQASQSQSQSFPPAPRGRGARGFGER
FGDQPRIFCLFCGNGKHTTRMCHVTQKQKEIAEAAQAQPKQIMHTASYHSPYI
PEYGNHDAVSASASQASWHQPPPPPLQQGQPEGQVAQHQRDFREQSEARTV
NSTVPESKHIV"
31090, .31758
31759, .31763
/note="sequence target duplication"
32800, .32855
/rpt_type=dispersed
36025, .37602
/note="pip(FIO)20725 RFLP marker"
join(<36233, .36579,37382, .>37925)
/product="hypothetical protein"
join(36233, .36579,37382, .37925)
/codon_start=1
/evidence=not_experimental
/product="hypothetical protein"
/protein_id="AAD20312.1"
/db_xref="GI:4416307"
/translation="WGSFKGHVLPGLTFLAVGLWRVWSSAARHAAEPSSFRVWSPA
PAPGPARULELVVAGGALADMCVEVLYSTHLRIIPAGGEVNPALNDLEHGMLLFF

```

```

LFGALASQLWPRHFLTDGALCLVAATAFTAEVLVFFPHSTHMGLEGYHYLLVL
LVGLCVAAVIGALLPDSFPADLASGVLVQGLWIFYOTALALYIGMILPAGCARDADA
HVECSRRAAQEAQLAISQLFALVFLALGYCAVAAAYGRPELPAVRHRRRA
AVELECRQDDAGAMECAI"
complement(38797, .38910)
/rpt_type=dispersed
complement(39147, .39151)
/note="target sequence duplication"
complement(39152, .49223)
/rpt_family="gypsy/Ty3-type retrotransposon"
/rpt_type=dispersed
complement(39152, .42406)
complement(42452, .45283)
/note="gypsy-type polypeptide"
/pseudo
/codon_start=1
complement(45973, .49223)
complement(49224, .49228)
/note="target sequence duplication"
complement(50243, .50571)
/note="similar to grande 1"
complement(50572, .60252)
/note="nested in grande 1 retroelement"
/rpt_family="copia-type retroelement PREM-2"
/rpt_type=dispersed
complement(50572, .50576)
/note="target sequence duplication"
complement(50577, .52000)
/note="copia-type retroelement PREM-2"
complement(<52949, .>56140)
/product="copia-type polypeptide"
/codon_start=1
complement(52949, .56140)
/product="copia-type polypeptide"
/protein_id="AAD20307.1"

Query Match 3.6%; Score 59.8; DB 8; Length 346296;
Best Local Similarity 62.3%; Pred.No.0.039; 92; Indels 5; Gaps 4;
Matches 160; Conservative 0; Mismatches 92; Indels 5; Gaps 4;

QY 942 AGTAGAGCACGTGGTGTATACGTTAAGAAACCCCTTTGTTACCGAGTGTAGTCCAAGG 1001
DB 206658 AATGAGGATTCGTCTGATACGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAAGG 206717

QY 1002 CTGCTAACT--GAATAATAGTAGAACACAGGACAGAGTGTAGAGAGAACTGATT 1059
DB 206718 CTCTTAATTAGAAATAGAAATCATAACACAGGACACAAATGTATAGGAGAACTGTT 206777

QY 1060 CTTTCTTACTATATGGTGTCTCTCCAAAGGTTACATGATATGGGATCTCTCTCTAT 1119
DB 206778 ATTCTTATTCATGTATT-CAGCTCATTTGTTACAAACAATGGTGTCTCTCTCTATAT 206836

QY 1120 TTATAGACAAACTAGGTTTCAGGCATATATGGCCACATAGGCCCTTCTTGGCCCAAGAAA 1179
DB 206837 ATAGGAT-AGAATTAGGTTTCGGCTAGATGATGATGATGATGATGATGATGATGAT 206894

QY 1180 GGTTCCTTAACACTACC 1196
DB 206895 GGTTCCTTAACACTCCC 206911

RESULT 6
PFMAL6P1_13
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMAL6P1 Accession AL844505
Fragment Name Begin End
PFMAL6P1_00 1 110000
PFMAL6P1_01 100001 210000
PFMAL6P1_02 200001 310000
PFMAL6P1_03 300001 410000
PFMAL6P1_04 400001 510000
PFMAL6P1_05 500001 610000
PFMAL6P1_06 600001 710000

```



```

/organism="Dermatobia hominis"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/strain="Ponta Grossa"
/db_xref="taxon:115427"
complement(<1..42)
/product="tRNA-Ile"
43..1611
1612..>1791
/product="12S ribosomal RNA"

ORIGIN
Query Match          3.4%; Score 56.2; DB 3; Length 1791;
Best Local Similarity 49.5%; Pred. No. 0.066;
Matches 145; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 15 CAATPACCAGTCATCTACATGATGATAAAATTTTATAGTTAGATTACTCTTGTGTAATTC 74
    |||||
Db 1242 CAATATTATATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1183

QY 75 AGAGATGAAAGAAATTCGCAAGAACCAATATTTTAATGATGATGCAATATACAA 134
    |||||
Db 1182 TAAAGAAATACATAAAATCAACCAATATTTTAAATTTTAAATTAATTAATTAATTTAT 1123

QY 135 ATTAAATTACACAATTATGTAAGATTACATTTGTTTACATAGAAATCAATTTCTAGA 194
    |||||
Db 1122 TTATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1063

QY 195 GTCATAATAATGCGCTAAATTAATTAATTTCTATCTATTTTCTTAAAGAAAAACACGCC 254
    |||||
Db 1062 AATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1003

QY 255 ATTAAGGACCATAGATGCGCTGCTCCATTTTATATATAGATAT 307
    |||||
Db 1002 AATATATATTGATTATATATATATATATATATATATATATATATATATATATATATAT 950

RESULT 9
AF151388/c
LOCUS
DEFINITION
Dermatobia hominis tRNA-Ile gene, partial sequence; and 12S
ribosomal RNA, partial sequence; mitochondrial genes for
mitochondrial products.
ACCESSION
AF151388
VERSION
AF151388.1 GI:7159046
KEYWORDS
mitochondrion Dermatobia hominis (human botfly)
SOURCE
Dermatobia hominis
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Oestridae; Dermatobia.
1 (sites)
REFERENCE
Lessinger,A.C. and Azeredo-Espin,A.M.
Evolution and structural organisation of mitochondrial DNA control
region of myiasis-causing flies
Med. Vet. Entomol. 14 (1), 71-80 (2000)
JOURNAL
2021065
MEDLINE
10759315
PUBMED
2 (bases 1 to 1792)
REFERENCE
Lessinger,A.C. and Azeredo-Espin,A.M.L.
Direct Submission
Submitted (14-MAY-1999) Centro de Biologia Molecular e Engenharia
Genetica - CBMEG, Universidade Estadual de Campinas - UNICAMP, Lab.
Genetica Animal, CBMEG, UNICAMP, P. O. Box 6109, Campinas, Sao
Paulo 13083-970, Brazil;
Location/Qualifiers
1..1792
/organism="Dermatobia hominis"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/strain="Alfenas"
/db_xref="taxon:115427"
complement(<1..42)

tRNA

```

```

/product="tRNA-Ile"
43..1612
1613..>1792
/product="12S ribosomal RNA"

ORIGIN
Query Match          3.4%; Score 56; DB 3; Length 1792;
Best Local Similarity 49.3%; Pred. No. 0.073;
Matches 146; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 12 TCACAATPACCAGTCATCTACATGATGATAAAATTTTATAGTTAGATTACTCTTGTGTAATCAT 71
    |||||
Db 1249 TCAATATTTTATATTCATTATATTAATTTTAAATATTTTAAATTTTAAATTTTAAATTTTAAAT 1190

QY 72 TTCAGAGCATGAAAGAAATTCGCAAGAACCAATATTTTAAATGATGATGCAATATA 131
    |||||
Db 1189 TAAATAAAATCATATAAAATCAACCAATATTTTATAATTTTAAATTAATTAATTTTATAATT 1130

QY 132 CAAATTTAAATCACAAATTAATGTAAGATTACATTTGTTTACATAGAAATCAATTTCT 191
    |||||
Db 1129 TATTATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1070

QY 192 AGAGTCATAATAATGCGCTAAATTAATTTCTATCTATTTTCTTAAAGAAAAACACAG 251
    |||||
Db 1069 TTAATTTAATAAATAAATTAATAATAAATTAAGTAATATATATATATATATATATATATATAT 1010

QY 252 CCATTAAGGACCATAGATGCGCTGCTCCATTTTATATATAGATAT 307
    |||||
Db 1009 ATAAATATATATTGATTATATATATATATATATATATATATATATATATATATATATATAT 954

RESULT 10
BV074228/c
LOCUS
DEFINITION
S212P6542FG4.T0 CZECHII/B1 Mus musculus STS genomic, sequence
tagged site.
ACCESSION
BV074228
VERSION
BV074228.1 GI:31190023
KEYWORDS
STS.
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 328)
REFERENCE
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)
JOURNAL
22354684
MEDLINE
12466852
PUBMED

CONTACT: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersti@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 328
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/SvImJ, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the WGSv3 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.
FEATURES
Location/Qualifiers
1..328
/organism="Mus musculus"
source

```

```

/mol_type="genomic DNA"
/strain="CZECHII/81"
/db_xref="taxon:10090"
/maps="+ 19 33-281 29021847-29022095"
/clone_lib="CZECHII/81"
<1..328

```

ST5 ORIGIN

```

Query Match          3.4%; Score 55.8; DB 11; Length 328;
Best Local Similarity 54.0%; Pred. No. 0.052;
Matches 114; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 37 ATAAATTTTATAGTTAGATTACTTCTGTGAATCATTCAGAGATCAAAAAAATCGCA 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 ATGAAGTTTAAATATAAATAATATAATATTAATATTTATTAATATAATATTAAT 213
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 97 AGAAAGCAAAATTTTAAATGAATGATCAATATACAAATTTTAATPACAAATTTATGTA 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 ATAATACAAATATATTAGCATAATATTAAATATAATATTAAAGTTTAATATAAATATA 153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 GATTACATGTTTGTCTTTCATAGAATCAATTTCTAGAGTCATATAATGCCCTAAATTA 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 GGTATAATGTAAAGTTTAAATATAATAGTTCAGCAAAACTGATTAGTCAATTAAGGCA 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 AATTCATCTATTCTTTTCCTTAAGAAAAA 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 AACGCATATGTTATTAGCGAAAAAACA 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 11
AC092635
LOCUS      AC092635      125316 bp      DNA      linear      PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-342E23 from 2, complete sequence.
ACCESSION AC092635 AC041020
VERSION    AC092635.2 GI:15638821
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 125316)
            Sulston, J.E. and Waterston, R.
            Toward a complete human genome sequence
            Genome Res. 8 (11), 1097-1108 (1998)
            99053792
            9847074
REFERENCE  2 (bases 1 to 125316)
            Trani, L., Cotton, M. and Spalding, L.
            The sequence of Homo sapiens BAC clone RP11-342E23
            Unpublished (2001)
REFERENCE  3 (bases 1 to 125316)
            Waterston, R.H.
            Direct Submission
            Submitted (19-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  4 (bases 1 to 125316)
            Waterston, R.H.
            Direct Submission
            Submitted (18-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE  5 (bases 1 to 125316)
            Waterston, R.
            Direct Submission
            Submitted (01-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
            On Sep 18, 2001 this sequence version replaced gi:14916220.
COMMENT    ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu/gsc
            Contact: sapiens@watson.wustl.edu

```

```

----- Summary Statistics
Center project name: H_NH0342E23
Drafting Center: WIBR
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-132H1; the clone sequenced to the right is RP11-375H3. 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-342E23; actual end is at base position 5825 of RP11-375H3.

Data from AC073062 was used to finish this clone, AC041020.

The sequence of AC041020 has been incorporated into AC092635.

FEATURES	source
repeat_region	1..125316
repeat_region	/organism="Homo sapiens"
repeat_region	/mol_type="genomic DNA"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="2"
repeat_region	/map="2"
repeat_region	/clone="RP11-342E23"
repeat_region	/clone_lib="RPCI-11"
repeat_region	43..137
repeat_region	/rpt_family="AcHobo"
repeat_region	397..636
repeat_region	/rpt_family="L1"
repeat_region	800..841
repeat_region	/rpt_family="AT_rich"
repeat_region	1476..1623
repeat_region	/rpt_family="MIR"
repeat_region	1943..2100
repeat_region	/rpt_family="MIR"
repeat_region	2207..2310
repeat_region	/rpt_family="L2"
repeat_region	2356..2404
repeat_region	/rpt_family="Mariner"
repeat_region	2447..2496
repeat_region	/rpt_family="Mariner"
repeat_region	2506..2649
repeat_region	/rpt_family="L1"
repeat_region	2916..2976
repeat_region	/rpt_family="ERV1"
repeat_region	2964..3035

```

repeat_region /rpt_family="(TA)n"
3058. .3162
repeat_region /rpt_family="(TA)n"
3207. .3230
repeat_region /rpt_family="AT_rich"
3215. .4207
repeat_region /rpt_family="L1"
4220. .4504
repeat_region /rpt_family="Alu"
4626. .4917
repeat_region /rpt_family="ERV1"
5260. .5425
repeat_region /rpt_family="L1"
5681. .5998
repeat_region /rpt_family="ERV1"
6019. .6191
repeat_region /rpt_family="(TAGA)n"
6226. .6264
misc_feature /rpt_family="ERV1"
6666. 6802
/notes="match to EST AA603518 (NID:G2437379) np07dl2.s1"
8592. .9336
repeat_region /rpt_family="L2"
9882. .9909
repeat_region /rpt_family="AT_rich"
10148. 10169
repeat_region /rpt_family="AT_rich"
10365. .10397
repeat_region /rpt_family="AT_rich"
11309. .11355
repeat_region /rpt_family="AT_rich"
11354. 11504
repeat_region /rpt_family="L1"
11500. .11576
repeat_region /rpt_family="L1"
11552. .11735
repeat_region /rpt_family="L1"
11865. 12101
repeat_region /rpt_family="L1"
11988. .12025
repeat_region /rpt_family="AT_rich"
12148. .12286
repeat_region /rpt_family="L1"
12354. 12476
repeat_region /rpt_family="L1"
12453. .12504
repeat_region /rpt_family="T-rich"
12477. .12784
repeat_region /rpt_family="Alu"
12785. .15212
repeat_region /rpt_family="L1"
14498. .14532
repeat_region /rpt_family="AT_rich"
15706. .15920
repeat_region /rpt_family="L1"
15925. .17405
repeat_region /rpt_family="L1"
16747. .16770
repeat_region /rpt_family="AT_rich"
17406. .17774
repeat_region /rpt_family="MaLR"
17775. .18148
repeat_region /rpt_family="L1"
18234. .18257
repeat_region /rpt_family="AT_rich"
18518. .18557
repeat_region /rpt_family="(CA)n"
18900. .18932
repeat_region /rpt_family="(TG)n"
18993. .19186
repeat_region /rpt_family="MER1_type"
20216. .20677
repeat_region /rpt_family="L1"

```

```

repeat_region 20996. 21268
/rpt_family="L1"
repeat_region 21307. 21979
/rpt_family="L1"
repeat_region 21590. 21633
/rpt_family="AT_rich"
repeat_region 21988. 22117
/rpt_family="L1"
repeat_region 22172. 22794
/rpt_family="L1"
repeat_region 22638. 22659
/rpt_family="AT_rich"

Query Match 3.4%; Score 55.8; DB 9; Length 125316;
Best Local Similarity 52.3%; Pred. No. 0.25;
Matches 123; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 32 ATGTGATAAATTTTATAGTTAGATTACTTCTTGTAATCAATTCAGAGAGTCAAAAAAAA 91
|||||
Db 94026 ATGTACTGAGTTTGTATTACATACATACAGCAATTAAATTTGATTAGAATAGCTTACAGGA 94085

QY 92 TCGCAGAAAGCAAAATATTTTAATGAATGCAATATACAAATTTAAATTAATACAAATTA 151
|||||
Db 94086 TAGAAAGAAATTTACTATATTACAAGATTTCATACAATCTCGAAGAGTGTGTAATATCA 94145

QY 152 TGTAAAGATTACATGTTTAGTTTCATAGAATCAATTTCTAGAGTCAATATAATGCTTAA 211
|||||
Db 94146 TTTGATAGTTCTTTATAGTAAGTTAAATACATTTATCTTTTAAATATATAAAATATATTA 94205

QY 212 ATTAATAATCTATTCTATTATTTTCTTTAAGAAAAAACAACCCATTAAAGGGACCA 266
|||||
Db 94206 ATAAAACTAATAATAACTTAACATTTCGAGATAGATGAATTCATTATAAAACCA 94260

RESULT 12
PFA929356 335050 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum strain 3D7, chromosome 9; segment 2/5.
DEFINITION AL929356 AL844508
ACCESSION AL929356.1 GI:233504968
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 335050)
AUTHORS Hall.N., Pain.A., Berriman.M., Churcher.C., Harris.B., Harris.D., Mungall.K., Bowman.S., Atkin.R., Baker.S., Barron.A., Brooks.K., Buckee.C.O., Burrows.C., Cherevach.I., Chillingworth.C., Chillingworth.T., Christodoulou.Z., Clark.L., Clark.R., Corton.C., Cronin.A., Davies.R., Davis.P., Dear.P., Dearden.F., Doggett.J., Feltwell.T., Goble.A., Goodhead.I., Gwilliam.R., Hamlin.N., Hance.Z., Harper.D., Hauser.H., Hornsby.T., Holroyd.S., Horrocks.P., Humphray.S., Jagels.K., James.K.D., Johnson.D., Kerhornou.A., Knights.A., Konfortov.B., Kyes.S., Larke.N., Lawson.D., Lennard.N., Line.A., Maddison.M., Mclean.J., Mooney.P., Moulé.S., Murphy.L., Oliver.K., Ormond.D., Price.C., Quail.M.A., Rabinowitz.E., Rajandream.M.A., Rutter.S., Rutherford.K.M., Sanders.M., Simmonds.M., Seeger.K., Sharp.S., Smith.R., Squares.R., Squares.S., Stevens.K., Taylor.K., Tivey.A., Unwin.L., Whitehead.S., Woodward.J., Sulston.J.E., Craig.A., Newbold.C. and Barrell.B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 22255708
PUBMED 12368867
REFERENCE 2 (bases 1 to 335050)
AUTHORS Pain.A., Hall.N., Bowman.S., Churcher.C., Quail.M. and Barrell.B.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT For more information about this sequence or the Malaria Project,

```

see http://www.sanger.ac.uk/Projects/P_falciparum.
Location/Qualifiers
1. .335050
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosomes="9"
complement(join(496. .1978,2292. .4594))
/gene="PF10330c"
complement(join(496. .1978,2292. .4594))
/gene="PF10330c"
/note="Hypothetical protein. Pfam match to PF00560,
Leucine Rich Repeat"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51752.1"
/db_xref="GI:23504973"
/db_xref="SPTREMBL:Q81382"
/translation="MRISKLFISGRQYKSCRSKSKKIFSPDPFVNTNFKIPKKK
WLFKTKQNNKYGYSRCSYNTDNGIENIEIERKRGNEKQSDNILLFDMHKKKE
KKRMLKNNHNYHHNNNNNTFNFRNQIIITVNLKNDIINDIYELVLPFDLDA
LSNVLNKSYGCGMNLQIYFKKLKRVHKKNLINLVTSKIIINNIYSLNLFYHTYK
KNFLHTSMYQFSFTLTKKLDLNVFTYQFISKEKKKAYIQLIKDDKIKDK
NTHVMNIIHMKYIHNLYFKIMNMFIRFVILYNKKEKKKNGFHIKENVMNVC
RTPLSNIIYFNKFKLKNFLINKRYIAFTGLRSQFNKLKSYKEKNNNININ
SINNVYIYINNIYNNLSNPDVHDHYSNKHAYQNIISFNLLKLHIFPKL
NKKNECKEILNINDMFIHTFNKRGKKKYPFLFYKKEIFNTNDFIIDRENV
KNPFYFHTLIGPTFPFNEKEKSGITCNPKKQKKKMIKKRLSYLYNK
TNPENISGFFGKRNKKIINLTNRKRVFNLNNELKYNSTNYIIFFLYDLIF
NHNDFLYKMYTNLEREEBEKTHSFFICTFNKKNKKSILKNILKDKLKNRI
ILFTNDKQLQSMCFKINIFYQNVNMLKYNHNNNNKNYDTPYHNNNINKKLIK
LYVDNKLMPGTINNKKNIKKVDFLQIKMNEENYVNDKQSPFIYFYEYEG
DDCFPNTIISFKKLKIYNEKDFEVNLNAEHLIKDVKIEGLORNIYNTBIIKKR
NNKRNKKKEVFDINNMENMNDINYNPKYDNCDDLKYNKERFDILLNELND
SNIKSYTLKEKSEGEKKKLVNMDLSNYPKESIDNLLSIMSKNIIHPKNIL
QLNRLTLNRKNLIIYPPYISNFDGLVNIHSHNYINGCNYIDKIDNMTYISKN
LCEEDIKNNINFGNEGWNKILPNLKNIYQNNYLOSFSLKLNKKHKNKYIN
ISFNKINPLDLFLKVNHEILNYSNTFMNVSNTSGHTKDINNMOBKNDIYHN
INKNETTQDBEITINDVHTIEQTSHTFCGQENKMLISDGTYNQHDNDNNVHN
KDNNEIIFHNLNLLKFFRPSLANLAIKTEVYQKFLYIKKEDYKYEQNYFDPK"
join(5068. .5415,6020. .6763)
/gene="PF10335w"
join(5068. .5415,6020. .6763)
/gene="PF10335w"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51753.1"
/db_xref="GI:23504970"
/db_xref="SPTREMBL:Q81381"
/translation="MYNDEIKILKEKYSFLKPYEQINQTFCIHLKCDIEIKAEAK
OKVVDIEDSEENKQIITDNTKHFLLSKYEKENILFYIKNHPNYFEILANTSY
HDIKFENIDFNKKTKETKLNKFPDSIKNIKKIVTHTGLYDNNMLVAYDENDV
NTQVFLSNDYCLNDVKNVNIYIOEKNKYNSDDIINVIDEQKLKLAHLKXTM
NDEQYDKRPNIRFLKLRKNKTYQNLISIQBENNNKINEHDPITYSINKPFDQ
ELYKIGKRCIKLGDGRKIFGVNKNQDIYVGVLDLDEPLNGSDGSKKKKLFCK
GDYGVGINISIVGVPFPPDINDLEEF"
complement(7546. .7638)
/note="possible exon"
complement(7844. .7891)
/note="potential exon"
complement(8133. .12401)
/gene="PF10340c"
complement(8133. .12401)
/gene="PF10340c"
/note="2 potential downstream exons possible
TMMW2.0 at aa 511-530 and 542-561"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51754.1"
/db_xref="GI:23504971"
/db_xref="SPTREMBL:Q81380"

JOURNAL Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 3 (bases 1 to 25132)
AUTHORS Hyman,R.W., Fung,B., Conway,A., Kurdi,O., Mao,J., Miranda,M., Nakao,B., Rowley,D., Tanaki,T., Wang,F. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
source
1..25132
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
136..182
/rpt_type=tandem
328..838
/rpt_type=tandem
1400..1425
/rpt_type=tandem
1556..1583
/rpt_type=tandem
1994..2022
/rpt_type=tandem
2249..2282
/rpt_type=tandem
2553..2591
/rpt_type=tandem
2885..2909
/rpt_type=tandem
2970..3091
/rpt_type=tandem
3136..3162
/rpt_type=tandem
3376..>4694
/locus_tag="PFL0565w"
join(<3376..3481,3717..3818,3990..4253,4432..>4694)
/locus_tag="PFL0565w"
join(3376..3481,3717..3818,3990..4253,4432..4694)
/locus_tag="PFL0565w"
/note="HMPfam hit to PF00226, DnaJ domain, score 2e-39"
/codon_start=1
/product="heat shock protein DnaJ homologue Pfj4"
/protein_id="AAN3202.1"
/db_xref="GI:23496646"
/translation="MSRRVNYEVLGPQADLTIVIKSKYRTLAMKHPDKPNNAE
ATERFKQISEAYELSDPKRRKYDLYGTENYADENDEFSNFHNFQFDAQRIPE
MFGDSSPFGNDSPFSDVMGSSFVDRGRVPRSDPDPFGSSFNVPFGSSPDNEM
DGGCFTSVETSTSGGFKNRVKTSTSKTSIINGKRVTRIVTKTLPNGTVERTV
TERBEDDRGNINIRQLPAHELRRNR"
3500..3544
/rpt_type=tandem
3873..3905
/rpt_type=tandem
3917..3956
/rpt_type=tandem
4817..4887
/rpt_type=tandem
4820..4874
/rpt_type=tandem
4936..4994
/rpt_type=tandem
5108..5144
/rpt_type=tandem
5319..5381
/rpt_type=tandem
5565..5636
/rpt_type=tandem
complement(<5840..>7003)
/locus_tag="PFL0570c"
complement(<5840..>7003)
/locus_tag="PFL0570c"
complement(5840..7003)
/locus_tag="PFL0570c"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36203.1"
/db_xref="GI:23496647"
/translation="MIVLLKNNLKNLTLPPLGNNKLVFMRGYNIEQTKGPIEKKT
TDNNEKMLNTLEDERVSKSKCNALFDECEELYRMKKVKDKESVEYDNPFIIEEL
NREELKITRNIFERINKKKCHVSNLSYKDIYDPYNNPFKVDNLFKVEYFNYSK
NASIVELKKORRAIKKSLKEQKLYNPYSSEYNNFNLMNDMDKEQMECHDKOEYK
WNSFNKKILNIKSPFTWRYTHLHPHENGILPRKINMTTKQOIQFKSICIA
RMALYPYDKKPEKDDLPLMDPLOFLDELHRYIQLNDLPAQALIKVMINRPSLN
YKVFYSKKEKDKNEKTAQNEQENVEKPEPSQMLQYKKNYENFYK"
7326..7363
/rpt_type=tandem
7629..7664
/rpt_type=tandem
7668..7702
/rpt_type=tandem
8021..8048
/rpt_type=tandem
8092..8133
/rpt_type=tandem
8274..8311
/rpt_type=tandem
8355..>17469
/locus_tag="PFL0575w"
join(<8355..17090,17320..>17469)
/locus_tag="PFL0575w"
join(8355..17090,17320..17469)
/locus_tag="PFL0575w"
/note="HMPfam hit to PF01593, Flavin containing amine
oxidase, score 1e-05"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36204.1"
/db_xref="GI:23496648"
/translation="WGTEKEDDEQDIVNRNDLKKKKNNKCDNMSNDISOSISNIS
DSISTNSNISNINISNINISNINISNINISNINISNINISNINISNINISNINIS
ISQNHDIKKVHIKNLIMTHKNINNYKNGNLTINYKKKGNDGENTKINTPRES
YDIFVYLSKSKSLKQSDMIKONMSQDFINVFQSKLKHNDLIMINKLYIYI
KFSODIDLHLPVFKREFFKTLIGNNILLAKYEVTEVLIKSLYKLTQKPTSN
FTIEDLVHNNPMKYEIYRIVINPFPIQNNKNGDHKKKINNINNKKKYIDHE
DINNVEYHOLEENQKTSIPFEKCYKNIKCMNFEKSHNMNKIOTVEDFOKENT
ENGKILYKKEKGINYNEKDMCMHSAFDNNMVQNNVLDNKNCKNTKKOKKN
NVIKPMVYINVDYNNKNNHNSLOENNYSLQIIPNNKQINNIRANVILNNEF
NLDVQFNTETWNKIESNKNDEERLQFEKEKQILKREKLRNCKRQKQOEK
ENKQEKIYKKEKLLILKEIKKLKEKELKEKEKKEKELKKEKELKKEKELKKEK
KROKEFLRKIKSEKKLKEKKLKEKELINSKEERKPELKSFKYSKKNHL
NKDKKTKQNKIDKSNQNNKNNNNKKDDINYLQFLTDQNNNTVDKKKKKTH
DISLSIKKKKCTKNLLNSINDIKSDLSFISTPTFNIEKEQEDRNEINDILKL
IEIKVWNTQEKNDYIECLHKSQKVNKKDKIKQQLLSKGNKILITKNGEHNQ
SDPLQNSDLNDDFLNNMTIYKKGKGTSDVMSVNDVSNVDISNVDPSNKD
ISNKLISYVDSKNNVHTNLSNIDNIHYNNKNNSENESEINELILKGLSKFG
DINKLSLOKILRIKESKGNKLETUVHIPPNVIERSAKKILVAKHEGIRKQGVKE
KYTKKLSLFGIVKPEDIKRGINNVTHTEDNQLYDVLIIAGISGLAASTYLN
KCNARFLVIEGRNIGRAFSTILPERIINNKLVPETVDLGANYLHCCDADLTNDK
KGTGKEFANCTEFNNEKMEKEDIDYEDFIWNKKKTKKKYKIVKKNKSYNN
KDNKNNKMGNSIYDQKEYVTNINLFEYKGLINDFAYLPYFVYNNLSQSEHN
QDDEKTKDNKISGGKNIVYLSNNKNIISCSNINYNNORVSDYDEYFSRRIKORP
KIKLRNNMLFKNNIDNIRNYFKYKKEKDNILNFKSLDELIHFVNNLSDCVHT
SSVLYPKNCTNKSKLILSNKOTRRRREYDKSVTQAEKLPKRVSLVCGKDNWE
STFYAHYNNENGKKISKFYIRVNLCDKIRVARMRNLFIFGSDNEYQDKIK
SEDOQKNNYVSDTTEKDSYFNNNDILKEKQIINNLIIDDTYHKEILFNNNHVK
NIKPNENNIIRNEFEASVLYNNKDKLENGKNAENGKNGENGENGNAENGKN
EONEQKNGKNGEENGKNGKNGKNGKNGKNGKNGKNGKNGKNGKNGKNGKNGK
KYDANNILSYDNHAIYNNYDGLNVKVVNVSSNTRNIDDEERKKEKEDISK
ELYNNKRRSRMDLLMECTEIEFTENNEENFTMEWMLVLTYSRYGSDLRE
TSIAMSRLPFSNLDIDMCPKYGNNYILKNITYDKIKNEQTPYISFPDLNISK
IYVQGWKLINLYSEEQNKIFLNTVAELVYIKDHQDKIYHLDNCOINQKQSSS
PLNTKKNNNNNNNNNNIESVNNNISDHYNDNNSFVPONEVSSKNKKYINNEF
LINNVIKDDYKLGENNEDDYKVMIKCKYDTFKKINKHGHGTDLNNCNYKNTIYA
KYIVVALPLGCLINDKKRKLKTCLEKPELHPLKLRALQNYKMGNNHKNILIRFPYN


```
FWPFDLSQINCIDQKQFELNLHAYKIGICILVHCPPWASCTGYGVIKKEHYIINECLY
TLNKMEKSGKLPILVDYIITKWQDDNFSFSGIAYPIYINCNDNDIYURSEHPINNP
RVVFCGEYLSKSYFQVGDGAYDTGIRAAEDIAHIGHLKKNNDTKNYKTVDYIYPQNK
PTNIPLPPIKFKLFGYITDGSDEALTIDYESSDENMNTVDINDINSNTNIDINN
NNSNTSNNNTKYNVCNNPIPISWMKKEYFLLYSQSQNIIFRLFKQWTKHNF
KODKNIHTKNETLADKCSYKEIKSNTVESFTNLSLNFEDKSNVDKSNISTLL
NLYEYNNKTDQIHFDRIKSDSYKSKYKNGSYLLELSKSESDNSNGDANNICTI
EYFNVDKKNKGHDINDVNDVNDKYSNINLETYSITNTHNFEXYNNISNLY
INKYTIKNDADFVYNEIEITILNKHNAEYKSYFVILLQTFNIFPKYSIDPTFIQ
KELFKIQAMKHILFENRNFQSLQDMNVLSQAHEKNDHIVSHNNHMDVDRKAYEQ
NTDIINEDDKETISYPERKMEVEVNNMKLSLNLIKSNMVGNNNDNDNNNGHIOKS
NSTDLELDKLNLSKNTLNKKNITNKDVMIOIOTNLINYNIDLOKNIKNIYNRNIC
VLNDLYKVLJHLISKNTLDQIOLSSNNKKNKSCGFLFELCVILOYLKHLKYDLFN
SVNTKVIQLLCDIYIYLLPFKHDILCYKNGGELNMCDSYNCNNSWHTYCLSSSE
HNNNNKDFWLPSCSNIDISKHVQRCYQNEIMENYWKRIYIYIKIFPFLSRTRK
INKRLDLKAHLRSVA"
repeat_region 8427..8510
/rpt_type=tandem
repeat_region 8428..8498
/rpt_type=tandem
repeat_region 8510..8652
/rpt_type=tandem
repeat_region 10256..10314
/rpt_type=tandem
repeat_region 11019..11081
/rpt_type=tandem
```

```
Query Match 3.3%; Score 55; DB 3; Length 253132;
Best Local Similarity 48.8%; Pred. No. 0.45;
Matches 148; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 7 ACGGCTCACAATACCGAGTCATCTACATGTCATATAATTTATAGTAAATTTATAGTAACTTCTGTA 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74935 ACTGTAATGAATCAATAATATTTATTTATTTATATACCTATTCCTATTTATTTATTTACTAC 74876

QY 67 ATCATTTTCAGAGGATGAATAAATAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATCA 126
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74875 GTGAAGTGAATAATGTAATAGAGATAATATATATATATATATATATTTTAAATGAC 74816

QY 127 ATATACAAATTTAATTACACAATTTATGTAAGATTACATGTTAGTTTCATAGAAATCAA 186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74815 ATATGAATAATAAATATTTAAGTATATAACATATAAATAATAATATATATATATATATA 74756

QY 187 TTTCTAGAGTCATAATATCGCTAAATTAATAATCTAATTTCTATTTTCTTCTTAAAGAAAA 246
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74755 TATATATATATATATATATATGATATATAATAATTTTAAATTTTAAATATATATATAAAT 74696

QY 247 AACAGCCCATTAAGGGACCATTAAGATCGCGCTGCTCCATTTTATATATATAGAGATA 306
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74695 TATAGAAAAATAAATAGAGATAAACAAGTACAGTTTCAGTATTGAAGGACACTA 74636

QY 307 TGA 309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74635 TGA 74633
```

```
RESULT 15
PPA929352/c 313050 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum strain 3D7, chromosome 5, segment 2/4.
DEFINITION AL929352 AL944504
ACCESSION AL929352.1 GI:23504570
VERSION AL929352.1
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1
AUTHORS Hall N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,
Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,
Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C.,
Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,
Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,
Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,
```

```
Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,
Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D.,
Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N.,
Lawson, D., Lennart, N., Line, A., Maddison, M., McLean, J., Mooney, P.,
Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,
Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,
Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,
Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,
Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and
Barrell, B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
Nature 419 (6906), 527-531 (2002)
22255708
12368867
2 (bases 1 to 313050)
Devlin, K., Baker, S., Davies, P., Mungal, K., Berriman, M., Pain, A.,
Hall, N., Bowman, S., Churcher, C., Quail, M. and Barrell, B.
Direct Submission
Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium,
The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
Location/Qualifiers
1..313050
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosomes="5"
complement (219..8039)
/gene="PFE0420c"
complement (219..8039)
/gene="PFE0420c"
/note="Similar to Plasmodium falciparum guanine
nucleotide exchange factor pfrccl (l327 aa) fasta scores:
E(): 9.6e-190, 97.459% id in 1338 aa"
/codon_start=1
/product="guanine nucleotide exchange factor, putative"
/protein_id="CAD51450.1"
/db_xref="GI:23504571"
/db_xref="SPTREMBL:Q81418"
/translation="MYEQDDEKDIHKYIFPPPKISTGLDHVACVAYYKPKSIFTW
GNSCNQLGLGINNRVFNVTVCPPDSFIVASICTNVATFVLKKNVNDIGCSVYS
FGKNGNLGYKKKRELPRFLEHEKNIQKVEQMLQKVNPLNMVLDAFGKKKE
KEKKKKKKKNDIKKEKEERINNEKNNNNNNINSEKKKKKNDINGHEAKYVQNK
KFFYNHNFKNIDIGLPSDDIYDDYINISGESLLEFADNISISSIOMEDTNNDEYVNN
IVNDHTNEQRQSDINETIDKSNDDNDYDDEKKKKKNIYKNEKLYSNIKKENDKHDW
FTPPMKIVFPKKNTKIKFTSCGMHTLAICTGGVLYSWGYNNGFGTNGTNGVYEP
TPIFLPHSVNDNNDNDDDDYDDNNNNNTYKPKKEDSIKWDKKNVYIHCAGSKH
SLACLHGLDIYSWYGGNGRLGLGNIKSYNKPOLIIGLRNKRILYVCAGTSHSGCIDS
NANYTWGNGKFKYLGHGNDIDIYYPKLVECLNHNKKVCMLSPGCFNSLALVKGDV
YWGTFNITNNYVYISKLPKQINTNYKICISIHASTYVCGITLVGDLISPGNKY
SYKEINDNDSDDLIKYLINKEVEHQYNGDNKSVHTKNNTIISYIVKKDEHDI
HYIKMRGMVYIKDVIYFNDLTMTYNNKINHIDNVTNNNDNYTDLNNGSHG
NIPKKNLVLKSKVLIIDGSDHPSVFLSGKVYTAGYNKNGELNGEFLKKKFGVP
ILLNICVNKLIKACGVYVLCVSDVGLVYWGKNDKQSLGIVIKDYPVPVKNLT
NVTNIFAGDHSACIINKLLDKGTNGVNNVNVNGVNGVNNYSDECGDIYNGNAE
SGKVGIGVDYDQTCILLPRKININKIYKCSLGNSSHSLFTNNELVCGSSNNKRDG
LFPKSMVCNIVSVPTKVLHNNIYIKDLAGNTYSIILSVDFGIYIFGFIKSGNG
NSNMCSEVLSYNETNVMYNGKYEHPFLTYDNKILFGTGNKNCQIILCDNKENNY
IKPKLITVFLKNNKIIESIYFSFHVNTFOFNNSBIFAMGYTNNVHLGIGIENIKY
LKHPKTIKTLWYDQKLDYNSCSDMELNEVNTNNMSYNNYIKYINBITLVKNI
NTTELQNETYCYGEEETENFIYIVIMNNLINCHLONLAKKEENNNMEYIKSF
EKDIDISYKHIEVLLNLIYEQKFNLYLNYONFLIUSNIANNKEQHLNLLPFSTY
IFQNRKLEQFIYIIQQQPVYLIILCLHNYKNVKNVQMLYMYKSNLNNKNTTHLY
HNNINSHIKNNYNIINSYINHTYEQNDALKYNSDKIQNKYSFYKNSNVLC
FIEDLVADPRNERVFTFLIKLGTIEELKNSLNVSIFINIDTSPFTLIQMLFEKK
EILINPAQCILIDNPNFVTLADOLSRVKKNIPLNQCVSNKMMQHTKHLQNT
PWSGAHRENNLLINDMYNDRVVDKNIHVKNSTHMDKNDGANKKKIKKKNNKK
NNRKKNNNNNNNNNNNGGDDYSGDDIYSGDDIYSGDDIYSGDDIYSGDDIYSGDDIY
EKDIFSGDDIYSGDDIYSGDDIYSGDGINHNSVKNLKNVKEPNISSEFIDNLND
EKDKEFLNFKGTQNESDKRFLVNDISCYKSGVFPYNIPIFNINNMENMYNNISQEI
```



```

ONENLONNRYVEINFCIPEELCKIPLNLKPPDMKLIILKELFKYFIIEKXNLDEI
NOTNKIYFQNHHTIYIPPKLILMALINPILKMNENIAHKSYPNIFPHILNIRKI
CDFLEILYKESLNYLNINISIKIIEHTFISLTVNINNINCNTHHIIYDAYTNL
CFYHLNRNSYVQLKLFQVCHIFNLFPRQYLMLSFNDFALDVNFVYTKNNIPIQ
KHGKHTLSTWGTWTKMDIKERERHLNIGLGDWGNIDRPGKGVSTFTGHDNKKI
NSMEGNEILLRAKADNLDNDRFMKGKMGINGVMKRGDPMGCGYPMGCGYPMGCGY
PMGNGYPMGCGYPMGCGYPMGCGYPMGCGYPMGCGYPMGCGYPMGCGYPMGCGY
AFNLPTIKITSDNDRKWKOKBLIPSEDDIEFFIKCLINYNKDLIRFLKKNWSI
CETKIPMQOYICRQKTLIENNEYIFSIIRKHYENKNNIYIISCEKNCPILEHCTD
TNLLILKSLKMNYSLOQKNNLILHLNKAIDIFLSDMTYVEFLDFEPANLSIY
KFQDILQKSMQEIYDFMLQNTYKSSQFQWNRNIALQISLNLKKGKNNYLLK
LYEKQRYIEDLIYOHKMKDITVLKNAIIFVSKLYIEKPIILLHSSYFKNLFFVEL
KKKKHTQFKLPYNSSTVHIYQMKDILNNQILNTHETLYTLVNVLTIEIPEDANNI
IKFSLILTKDQNKNIINEHTFTGSEIYSYMNSSPYIILYPPFKYSKTYLCSITGLHFMH
LJLNFVIDLY"
9222..10010
/gene="PFE0425w"
9222..10010
/gene="PFE0425w"
/gene="PFE0425w"
/note="Weakly similar to Arabidopsis thaliana F8n16.20
protein F8n16.20 TR:Q9ZV23 (EMBL:AC005727) (312 aa) fasta
scores: E(): 2e-12, 28.763% id in 299 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51451.1"
/db_xref="GI:23504572"
/db_xref="SPTREMBL:Q81417"
/translation="MPSKGRVMPADNRLPVSASLKTSEIWNKSVGVDPYASYEIN
KKKEEDINEKAKNLFSLRITGITSTTIFGACTVCHNIGHLPYQCRNFISLEKFD
INNWNNEEDDIKERKNLGLMSDDGSDSHSDSHSDSHSDSHSDSNDSNSDSNS
HSDSNHSDSYKXKXSHHSHKXKXSHHSHKXKXSHHSHKXKXSHHSHKXKXSHHSHK
SDEKHKYISSKKKKIKKEEKKXKXSHHSHKXKXSHHSHKXKXSHHSHKXKXSHHSHK
11261..15733
/gene="PFE0430w"
11261..15733
/gene="PFE0430w"
/note="Similar to Rattus norvegicus RNA helicase TR:Q62780
(EMBL:U25746) (1032 aa) fasta scores: E(): 2.4e-59,
38.387% id in 719 aa"
/codon_start=1
/product="ATP-dependent RNA helicase, putative"
/protein_id="CAD51452.1"
/db_xref="GI:23504573"
/db_xref="GOA:Q81416"
/db_xref="SPTREMBL:Q81416"
/translation="MNIINFYIKSKKEKKKKDKMDIKVKKQDHDDSDVLKRDN
ISSKKKQVERNYIDNSYNEEEKKKKKKDKSKKSYDDSYNKTKNKIYNDHKITDE
YPSYAKKSSYHVHREKYPDEKRYEHIKKKRSYDYSNYDESDFSSTSY
YNERNKYYKVEDKKKEKEKEKERKRYHHHTDDDDHLESYSSRSSRSIN
IKEKQKHKLKYSHEKERKRYNSANSYSSYDEKTEKDKKKKSYHHDVFLDDKKN
VHNYESDRETKDESNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QEI CENKKEIQKVENSNVKITENNLSRLERLKFARKIKTNDTITNEKPKFT
LNTKYNKLNLI FIDDOEDSDNKEFEFDIKENKINQVNDNNHKKDDTKKIE
NHKIKEDSNNTDHWIEINNKNEDALDLFMKEIEQACEEKQNVKKSITLDEIYTD
MKNQYEPFINPKNNEEKIYNAKVGNNIQQGNHEKQDSNVNTQEDLKKNENITK
NTNNYNSDNNNNDDDDDDDDVFKHFLVEELKKKTEEDMKENKNNWNESSLTD
SKNKKNGKSGDASFSSDDTDKLSNKSSELNENYENGMKKNNKLLLVNHDIDIFI
KXNIVQVEITNMKSDVDFPKNGNLIIVRGKNCPRPVQFYQCGPLSKILQILEK
KNFKWMYINQWOTIPALMGDRVIAIAETGSKTSLYFPIRVIRYHQLFELNNDGPI
SIILPTRELISQVNEAKYICKAVNIBILAVGGSNTARQLKVLKGVIELVGPGR
IIDLITISCKVTNLRVSFVLEADRLDLGFFSOIYNILNRCKDKQKAMISATF
PNYIQMAKKLLYKPIELIIVGEGKTNNIYQFVEIIBESKVFRLLKLGWIKYGL
VLI FVYKQIEADLLVLELYKDYNNLLVLLGGQDQDTRQFTLEKFKENKNTIATSVM
ARGDILKNILVINTQCPDIEDYIHRIGRTGRSNNIGYATFIFILPNTKAYDINL
LKNNIYLYNKTLIDIPDLENMIIEYTKINSINEKQKNNKNGLYKKGKGYKFTPDKSR
HQIDIALAKELQENDENEKQLQVDDELGNIIEHNNNTNVCNLKNDQNKYDFE
DYKESNIYPTIHNNSNMFDKNGISKNMSKKNTEKKNTOYTDITKVEQFIQR
IKQNDTMLSIAKKNINELMKTYNFLSQENHDKFLKNAQYEEFPKIAEKALKYTEH
IKDEKEXLIFNKTKEDIKKLIQNKIQSDTNESIQRNLLSALRTKMKKISYPLP
HYITLEDNNLIEEFTYINPQHRUKISHKOVAKIADMGSATCQIKGYSNPSQPNK
QNFLDITKQLHIEILAPTYNQIARNELNLSLNNFMISCNIKAKRASTGWAPP"
complement (16588..17442)
/gene="PFE0435c"

```

CDS

```

complement (16588..17442)
/gene="PFE0435c"
/note="Similar to Rickettsia prowazekii single-strand
binding protein ssb or rp836 SW:SSB_RICPR (Q92CC2) (152
aa) fasta scores: E(): 1.1e-16, 38.158% id in 152 aa
Signal peptide predicted by SignalP 2.0 HMM (Signal
peptide probability 0.896, signal anchor probability 0.017)
with cleavage site probability 0.471 between residues 25
and 26"
/codon_start=1
/product="single-strand binding protein, putative"
/protein_id="CAD51453.1"
/db_xref="GI:23504574"
/db_xref="SPTREMBL:Q81415"
/translation="MGKRLCIVFPLLIYFNVVLHRTYGIIGDVKXHOLNNIINKRI
IKSRKISMEFKINLQDNFENKRFYNNMNRNMVNEKSLNKLIMLIGRVGCEPDIKILG
GDVATFSLATNEFMRDNRNTNELKSKTDWHRIVVVDONIVDLIDKVLKRRVYVQGS
LHTRKWHHTNDMSQPKQITTEILSYNKGDLIFLDDKRNQRRNNNINSENQOHLNN
EHINNINNNNGNDFMLQNSDKLIIEDEFTDLDDNNEENNFTQSNSETFDKQGIYDK
MNVQBFEE"
20022..29426
/gene="PFE0440w"
20022..29426
/gene="PFE0440w"
/note="Weak similarity to Plasmodium falciparum Pfc0425w
protein pfc0425w TR:O77336 (EMBL:Z99547) (4550 aa) fasta
scores: E(): 7.2e-23, 22.792% id in 3238 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD51454.1"
/db_xref="GI:23504575"
/db_xref="SPTREMBL:Q81414"
/translation="MSTCNNIVDLVRNKKELLTGKKMRKSKSNIDRNTIDNENKLL
DKNIIILNFLNSEKLIKETQLNINISLRSATYLLKIEKIKLISCNWKOVPISYIK
KSAENILLKKNVEYIKNLTKLNKAQNNINNDINIDNIYNSCVNLVNNKDEKK
FSKHLNSSSSSLSSYSLNPSDDSCFNSKQNNPFNSIKNNSTFTGNAQTKNK
LDNHTSKTKIKNDKIYDKIKDINKDICDKYDKYDKYDKYDKYDKYDKYDKYDKYD
KIYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYDKYD
YSNKIINYKGNFYDNNKNEIKNNEYIYDDQDLDSVNGLCVTNGKNNIFCFLNPQE

```

gene

CDS

Query Match

Best Local Similarity 3.3%; Score 55; DB 3; Length 313050;

Matches 109; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY	26	ATCTACATGTGATAAATTTTATAGTTAGTACTTCTTGTATCATTTTCAGAGGATGAA	85
DB	199656	ATATATATATATATATATATATATATATATATATATATATATATATATATATAT	199597
QY	86	AAAAATCCACAGAAGCAATATTTTAAATGAATGATGCAATATACAAATTTAATTA	145
DB	199596	AAAAAAAAAAAAAAAAAATTTTTTTTAAATTTTAAATCATTAATAAAAAAAAAA	199537
QY	146	CAATATGTAAGATTACATTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAAT	205
DB	199536	AAAAATTAAGTTTATATCAATTTGTACATATTAATGTGAACATACAAAAAATGT	199477
QY	206	GCCTAAATTAATAATCTAT	224
DB	199476	GCATAAAAAAAAAAATAT	199458

Search completed: September 2, 2004, 09:38:04
Job time : 6308 secs

gene

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 06:07:12 ; Search time 639 seconds

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcacggtcacatac.....ctcgtagtttgggacggcg 1658

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68.8	4.1	2000	7	ADA71938 Rice gene
2	53.2	3.2	7498	6	ABL32257 Human imm
3	52.2	3.1	6012	6	ABL31371 Signal tr
4	52.2	3.1	6012	6	ABL70328 Chemical
5	52.2	3.1	6012	6	Aas61275 Human gen
6	51.4	3.1	6265	2	Aax08523 NBP46 (ro
7	50.8	3.1	12763	6	ABL32303 Human imm
8	50.8	3.1	113515	6	ABL34174 Human imm
9	50.6	3.1	17280	4	AAS46772
10	50.4	3.0	1134	6	ABT07549
11	50.4	3.0	15767	6	ABL33206 Human bre
12	50.4	3.0	15767	6	ABL33206 Human imm
13	50.2	3.0	8056	7	ABL34552 Human met
14	50.2	3.0	17142	4	ABZ10246 Haematopo
15	50.2	3.0	17142	4	AAS45498 Chemical
16	50.2	3.0	17142	6	ABL34107 Human imm
17	50	3.0	8067	6	ABK28430 DNA trans
18	49.4	3.0	14551	6	ABN80224 Human che
19	49.2	3.0	1501	7	ABL34584 Human met
20	49.2	3.0	1501	7	ABZ10042 Haematopo
21	49	3.0	2000	7	ADE84086 Human lym
22	48.8	2.9	8961	6	ADA71938 Rice gene
23	48.8	2.9	8961	6	ABK28428 DNA trans
					AB149380 Human pol

C	24	48.6	2.9	5991	6	AAS61087
	25	48.4	2.9	632	5	ABA06279 Soy bean
	26	48.4	2.9	736	5	ABA06275 Soy bean
	27	48.4	2.9	4590	1	AAN60472 Sequence
C	28	48.2	2.9	5313	6	ABL32763 Human imm
C	29	48.2	2.9	5768	6	ABL34163 Human imm
C	30	48.2	2.9	7276	6	ABL32902 Human imm
C	31	48.2	2.9	13038	2	ABL33275
C	32	48	2.9	2418	6	AQ27886
C	33	48	2.9	6053	4	AAS45371 Chemical
C	34	48	2.9	6053	4	ABK28208 DNA trans
C	35	48	2.9	7657	4	AAS45477
C	36	48	2.9	7657	4	ABK28227
C	37	48	2.9	33053	6	ABQ67006
C	38	47.6	2.9	5499	6	ABQ66972
C	39	47.6	2.9	9963	6	ABL32694
C	40	47.6	2.9	18997	6	ABL32570 Human imm
C	41	47.6	2.9	18997	6	ABK33948
C	42	47.6	2.9	18997	7	ADA20352
C	43	47.6	2.9	18997	7	ADA84159
C	44	47.6	2.9	113515	6	ABL34175
C	45	47.4	2.9	6078	6	ABK28227

ALIGNMENTS

RESULT 1
ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y, Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G; WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to


```

FT intron /number= 6
FT 4386..4694
FT /tag= 1
FT /number= 6
FT 4695..4909
FT /tag= m
FT /number= 7
FT 4910..5014
FT /tag= n
FT /number= 7
FT 5015..5198
FT /tag= o
FT /number= 8
FT 5199..5300
FT /tag= p
FT /number= 8
FT 5301..5610
FT /tag= q
FT /number= 9
XX WO9907223-Al.
XX 18-FEB-1999.
XX 05-AUG-1998; 98WO-US016261.
XX 06-AUG-1997; 97US-00907226.
XX (REGC ) UNIV CALIFORNIA.
XX Etzler ME, Murphy JB;
XX WPI: 1999-167136/14.
XX P-PSDB; AAW85684.
XX New polynucleotides encoding Nod factor binding lectins - useful for
XX production of transgenic plants which are able to fix nitrogen.
XX Disclosure; Page 44-47; 57pp; English.
XX The NBP46 root lectin is instrumental in recognising and binding to
XX nitrogen fixing rhizobial bacteria via a lectin-carbohydrate interaction.
XX The production of transgenic plants comprising an expression cassette
XX expressing the NBP46 root lectin is advantageous since it would mean that
XX non-leguminous plants could fix nitrogen from the atmosphere, lessening
XX the need for the addition of nitrogen containing fertilizer to soil. This
XX would lead to higher crop yields where soil has been overplanted and
XX replenishment of the depleted soil with usable nitrogen. Alternatively,
XX expression of NBP46 can be used to modulate oligosaccharide signalling in
XX the plant. The nucleic acid sequences can be used to inhibit expression
XX of an endogenous gene and also to suppress endogenous NBP46 gene
XX expression
XX
XX Query Match 3.1%; Score 51.4; DB 2; Length 6265;
XX Best Local Similarity 49.4%; Pred. NO. 0.037;
XX Matches 133; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
XX
XX 24 TCATCTACATGTAATAATTTTATAGTAGATTACTTCTGTGAATCATTTTCAGAGCATGA 83
XX |||||
XX 3725 TCCGCTTAAGGTAAATAATTGCGAGAAAAAATACATATGTAATTAATATATATAT 3666
XX |||||
XX 84 AAAAAAATCGCAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATA 143
XX |||||
XX 3665 AATATAACATAAAATTTAATATTTTATTTATTTATTTATTTATTTAATAAATAAATA 3606
XX |||||
XX 144 CACAATTTATGAATTTACATTTTGTAGTTTCATAGAATCAATTTCTAGAGTCATAATA 203
XX |||||
XX 3605 AATATAATATTTAATTAATTTATTTTATTTAATTAATAAATAAATAAATAAATA 3546
XX |||||
XX 204 ATGCCATAATTAATTTCTATTCTATTTTCTTAAAGAAAAAAGCCGCAATTAAGGGA 263
XX |||||

```

```

Db 3545 AAATATAAATTTAAATAAATAATTTTATTTTTCATATAATAATAATAATAATATT 3486
QY 264 CCATTAGATGCGCGCTGCTCCATTTTT 292
Db 3485 AAAATATAAATAAATAATTTTCACTTTTT 3457

RESULT 7
ABL32303/c
ID ABL32303 standard; DNA; 12763 BP.
XX
AC ABL32303;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 276.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cystostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 276; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
XX Sequence 12763 BP; 3481 A; 219 C; 3026 G; 6037 T; 0 U; 0 Other;
XX
Query Match 3.1%; Score 50.8; DB 6; Length 12763;
XX Best Local Similarity 56.6%; Pred. NO. 0.072;
XX Matches 94; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
XX
QY 83 AAAAAAATCGCAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATT 142
XX |||||
XX 2935 AAAAAAATTTTAAATTTAAACACCATTAATAATAATAATAATAAATTTTATT 2876
XX |||||
QY 143 ACACAATTTATGAGATTACATTTGTTAGTTTCTAGAAATCAATTTCTAGAGTCATAAT 202
XX |||||
Db 2875 TATTATTATATTTTAAATAATACCTTTTAAACAAATTTAAATAATTTCTTCTATTCT 2816
XX |||||
QY 203 AATGCCTAATAATAAATTTCTATTCTATTTTCTTCTTAAGAAAAA 248
XX |||||

```

Qy

205 TGCCATAAATTAAAAATTCATTTCTATTTTTCTTAAAGAAAAA 246

Dp

11282 TACAATAAATAATATAAAAATTTCTTTTAAATCAAAATTCAAAA 11241

RESULT 9
AAS46772/c
ID AAS46772 standard; DNA; 17280 BP.

XX AAS46772;

XX
DT 18-DEC-2001 (first entry)

XX Tumour suppressor gene derived chemically modified sequence #496.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.

PD 20-SEP-2001.

PF 15-MAR-2001; 2001WO-EP002955.

PR 15-MAR-2000; 2000DE-01013847.

PR 07-APR-2000; 2000DE-01019173.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

XX

PT genes and oncogenes, useful in designing primers and probes for analyzing PT diseases associated with cytosine methylation state e.g. cancer.

XX

bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Ss). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or

Sequence 113515 BP: 31803 A: 1174 C: 24020 G: 56518 T: 0 U: 0 Other: 0

Query Match	3.1%	Score 50.8	DB 6	Length 113515
-------------	------	------------	------	---------------

Matches 115; Conservative 0; Mismatches 107; Indels 0;

QY 25 CATCTACATGTGATAAATTTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGGATGAA 84

Db 11462 CAAATATATATACAAATATATATATACAAATATATATTTAAATATACTACTACATAAA 11403

QY 85 AAAAAATCGCAAGAAAGCAAATATTTTAAATGAATGATGCAATATACAAATTTAATTAC 144

Db 11402 AAAATAACAAATTAAATAATATTTTCTATTATTTTAAATATTTCTATTCTATTACAATCAC 11343

QY 145 ACAATTATGTAAGATTACATTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAATAA 204

Db
11342 AATAAAATAAAATTTTAAATCATATTCAAAAAAATACCATATTAAAAACACAAAAAAT 11283

[illegible]

RESULT 10
ABT07549
ID ABT07549 standard; cDNA; 1134 BP.
XX
XX AC ABT07549;
XX
XX 14-NOV-2002 (first entry)
XX
XX DE Human breast cancer associated coding sequence SEQ ID NO: 28.
XX
XX Human; breast specific gene; breast specific protein; breast cancer;
KW gene therapy; cytostatic; gene; ss.
XX
XX OS Homo sapiens.
XX
XX WO200264611-A1.
XX
XX 22-AUG-2002.
XX
XX 12-FEB-2002; 2002WO-US004197.
XX
XX 13-FEB-2001; 2001US-0268292P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;
PI Sun Y, Liu C;
PI
XX
XX WPI; 2002-657582/70.
XX
XX New breast specific nucleic acids and proteins, useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
PT non-cancerous disease states in breast tissue, and in gene therapy.
XX
XX Claim 1; Page 171; 367pp; English.
XX
XX The present invention provides human breast specific coding sequences and
CC proteins. These can be used in the diagnosis and treatment of breast
CC cancer and non-cancerous diseases of the breast. The present sequence is
CC a coding sequence of the invention
XX
XX Sequence 1134 BP; 418 A; 141 C; 170 G; 405 T; 0 U; 0 Other; 0 Q

	Query Match	3.0%;	Score 50.4;	DB 6;	Length 1134;
	Best Local Similarity	51.3%;	Pred. No. 0.032;		
	Matches 117;	Conservative 0;	Mismatches 111;	Indels 0;	Gaps 0
QY	11	CTCACAATACCAGTCATCTACATGTGATAAATTTTATAGTTAGTACATCTCTTGTAATCA	70		
Db	29	CTTGAAGTTCTAGTCAATTTAATTGTGTCCAATAAGTTTTCGAAATCTCCTTTTTAACT	88		
QY	71	TTTCAGAGGATCGAAAAAATATTCGCAAGACCAATATTTTAAATGAATGATGCAATAT	130		

Db	89	TCCAGAAATTC	TATTATAATAGTGTACATTTTACCAATGCCATGTATAGCAACAG	148
Qy	131	ACAAATTTAA	TACACAAATATGAAGATACATGTTTAGTTTCATAGAAATCAATTC	190
Db	149	ACACCTTTT	TAGAAAGGATAGTATCATCAATTTGTTTTTAAAAAACAATTC	208
Qy	191	TAGAGTCATA	TAAATGCGCTAAATTAATAATCTATTCTTTTCCTTA	238
Db	209	CAGACTACTA	ATTTGGCATAGAAATATCTTTTAAAAATGCAACATA	256

RESULT 11
ABL33206
ID ABL33206 standard; DNA; 15767 BP.

XX
AC ABL33206:

XX
DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1179.

Human; immune system disease; cytosine methylation; antiasthmatic;
antiarteriosclerotic; antianemic; cyrostatic; nootropic;
neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
antirheumatic; antithratic; antidiabetic; antipsoriatic;
antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
ds.

XX Homo sapiens.
OS

XX
PN WO200200928-A2.

XX
PD 03-JAN-2002.

XX PF 02-JUL-2001;

30-JUN-2000; 2000DE-01032529.

FR 01-SEP-2000; 2000DE-01043826.
XX
PA (PRIC-) PRICENOMICS NC

PI Olek A. Pienenbrock C

XX
DR WPI; 2002-130909/17.

XX Nucleic acid comprising

Pr diagnosis and treatment of diseases associated with abnormal cytosine methylation.

XX
PS
YY

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines in the promoter

can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and other diseases of the eye.

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative
CC diseases. The present sequence is a gene of the invention
CC

XX
SQ Sequence 15767 BP; 4558 A; 123 C; 3384 G; 7702 T; 0 U; 0 Other;

	Query Match	3.0%;	Score 50.4;	DB 6;	Length 15767;
Best Local Similarity	48.3%	Dead No.	0	000	

Best local similarity 49.3%; Pled. No. 0.099;
Matches 132; Conservative 0; Mismatches 136; Indels 0; Gaps 0

33 TGTGATAAATTTATAGTTAGATTACTTCTGTGAATCATTTCAGAGGATGAAAAAAAAAT 92

QY	93	CGCAGAAAAGCAATATTTTAAATGAAATGATCAATATACAAATTTAAATTCACAAATTAT	152
Db	14988	TATTTTGAGATGTGAAGTTAAATTTTGAAGAGTTTAAAGTTTAAATTTATTTTTT	1504

	15048	15049	15050	15051	15052	15053	15054	15055	15056	15057	15058	15059	15060	15061	15062	15063	15064	15065	15066	15067	15068	15069	15070	15071	15072	15073	15074	15075	15076	15077	15078	15079	15080	15081	15082	15083	15084	15085	15086	15087	15088	15089	15090	15091	15092	15093	15094	15095	15096	15097	15098	15099	15100	15101	15102	15103	15104	15105	15106	15107	15108	15109	15110	15111	15112	15113	15114	15115	15116	15117	15118	15119	15120	15121	15122	15123	15124	15125	15126	15127	15128	15129	15130	15131	15132	15133	15134	15135	15136	15137	15138	15139	15140	15141	15142	15143	15144	15145	15146	15147	15148	15149	15150	15151	15152	15153	15154	15155	15156	15157	15158	15159	15160	15161	15162	15163	15164	15165	15166	15167	15168	15169	15170	15171	15172	15173	15174	15175	15176	15177	15178	15179	15180	15181	15182	15183	15184	15185	15186	15187	15188	15189	15190	15191	15192	15193	15194	15195	15196	15197	15198	15199	15200	15201	15202	15203	15204	15205	15206	15207	15208	15209	15210	15211	15212	15213	15214	15215	15216	15217	15218	15219	15220	15221	15222	15223	15224	15225	15226	15227	15228	15229	15230	15231	15232	15233	15234	15235	15236	15237	15238	15239	15240	15241	15242	15243	15244	15245	15246	15247	15248	15249	15250	15251	15252	15253	15254	15255	15256	15257	15258	15259	15260	15261	15262	15263	15264	15265	15266	15267	15268	15269	15270	15271	15272	15273	15274	15275	15276	15277	15278	15279	15280	15281	15282	15283	15284	15285	15286	15287	15288	15289	15290	15291	15292	15293	15294	15295	15296	15297	15298	15299	15300	15301	15302	15303	15304	15305	15306	15307	15308	15309	15310	15311	15312	15313	15314	15315	15316	15317	15318	15319	15320	15321	15322	15323	15324	15325	15326	15327	15328	15329	15330	15331	15332	15333	15334	15335	15336	15337	15338	15339	15340	15341	15342	15343	15344	15345	15346	15347	15348	15349	15350	15351	15352	15353	15354	15355	15356	15357	15358	15359	15360	15361	15362	15363	15364	15365	15366	15367	15368	15369	15370	15371	15372	15373	15374	15375	15376	15377	15378	15379	15380	15381	15382	15383	15384	15385	15386	15387	15388	15389	15390	15391	15392	15393	15394	15395	15396	15397	15398	15399	15400	15401	15402	15403	15404	15405	15406	15407	15408	15409	15410	15411	15412	15413	15414	15415	15416	15417	15418
--	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

RESULT 12
 ABL34552
 ID ABL34552 standard; DNA; 15767 BP.
 XX AC
 XX ABL34552;
 DT 26-MAR-2002 (first entry)
 XX DE
 XX Human metastasis associated gene SEQ ID NO: 105.
 XX DE
 XX Metastasis associated gene; cytostatic; gene therapy; cancer;
 KW cytosine methylation; gene; ds.
 XX OS
 XX Homo sapiens.
 XX OS
 XX WO200177376-A2.
 XX EN
 XX 18-OCT-2001.
 XX PD
 XX 06-APR-2001; 2001WO-EP003970.
 XX PF
 XX 06-APR-2000; 2000DE-01019058.
 XX PR
 XX 07-APR-2000; 2000DE-01019173.
 XX FR
 XX 30-JUN-2000; 2000DE-01032529.
 XX FR
 XX 01-SEP-2000; 2000DE-01043826.
 XX FR
 XX (EPIG-) EPIGENOMICS AG.
 XX FA
 XX Olek A, Piepenbrock C, Berlin K;
 XX PI
 XX WPI; 2002-010922/01.
 XX DR
 XX New nucleic acid derived from chemically treated metastasis genes, useful
 PT for diagnosis of cancers by analysis of cytosine methylation, also for
 PT treatment.
 XX PT
 XX Claim 1; SEQ ID NO 105; 23pp + Sequence Listing; English.
 XX PS
 XX The present invention provides a number of human metastasis associated
 CC genes which are modified by cytosine methylation. The sequences can be
 CC used in the diagnosis and treatment of cancer. The present sequence is
 CC one of the genes of the invention
 XX CC
 XX sequence 15767 BP; 4558 A; 123 C; 3384 G; 7702 T; 0 U; 0 Other;

	Query Match	3.0%;	Score 50.4;	DB 6;	Length 15767;
	Best Local Similarity	49.3%;	pred. No. 0.099;		
	Matches 132;	Conservative 0;	Mismatches 136;	Indels 0;	Gaps 0;
Qy	33	TGTGATAAATTTTACGTAGATTACCTCTTGTAATCATTTTCAGAGGATGAAAAAATAAT	92		
Db	14988	TAITTTTGAGATGCTGAAGTTAAATTTTAAAAATTTTGATGAGTTAAAGTTTAAATTTATTTTT	15047		
Qy	93	CGCAAGAAAGCAAAATATTTTAAATGAAATGCAATATACAAATTTAAATTCACAAATTAT	152		
Db	15048	TGSTAAAGAAATATTTTATAGATAAATGTTGTGATATTTTATATGTTTATATACGAAAGTAT	15107		

Qy	153	GT	RAAGATTACATGTTAGTTTCTTCATGAAGAAATCCAAATTCCTAGAGTCATATAATATGCCTAAA	212
Db	15108	GT	AATCTTTGAATGTTTTATTTTATTTTGGTGATGTTTATTTAAGTCGATTTGTTTGTGTTTAAA	15167
Qy	213	TT	AAAAATCTATCTATTTCTATTTTCTTAAAGAAAAAAAACAGCCCAATTAAGGACCATATAGA	272
Db	15168	TT	GAAATTAATAATATTTTATTTTAAATTAATAATTAATAATTAAGTGTATTTAGAG	15227
Qy	273	TG	CGCGCTGCTCCATTTTTTATATATA	300
Db	15228	AA	CTTTTATATATTTATTTTGTGTTTTTTTAA	15255

RESULT 13	
ABZ10246	
ID	ABZ10246 standard; DNA; 8056 BP.
XX	
XX	
AC	ABZ10246;
XX	
DT	16-JAN-2003 (first entry)
XX	
DE	Haematopoietic cell proliferation disorder related DNA sequence #386.
XX	
XX	Human; haematopoietic cell proliferation disorder; cytostatic;
KW	gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW	cytosine methylation state; gene; ds.
XX	
OS	Homo sapiens.
XX	
XX	WC200277272-A2.
PD	03-OCT-2002.
XX	
PF	26-MAR-2002; 2002WO-EF003401.
XX	
PR	26-MAR-2001; 2001US-0278333P.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI	Olek A, Pispembrock C, Adorjan P, Grabs G, Iesche R, Leu E;
PI	Levin A, Lipscher B, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI	Schwowe I, Ziebarth H;
XX	
DR	WPI; 2003-018942/01.
XX	
PT	Detecting and differentiating between hematopoietic cell proliferative
PT	disorders, comprises contacting a target nucleic acid with a reagent that
PT	distinguishes between methylated and non-methylated CpG dinucleotides.
PS	Claim 28: SEQ ID NO 386: 117bp; English.
PS	

The present invention describes a method for detecting and differentiating between haematopoietic cell proliferative disorders associated with at least 1 gene and/or their regulatory regions in a subject. The method comprises contacting a target nucleic acid in a biological sample obtained from the subject with at least 1 reagent, which distinguishes between methylated and non-methylated CpG dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118 represent specifically claimed nucleotide sequences from the present invention. Oligonucleotides from the present invention can be used: for differentiating between healthy haematopoietic cells and proliferative disorder haematopoietic cells; for differentiating between acute lymphocytic leukaemia and acute myelogenous leukaemia; as probes for determining the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) of haematopoietic cell proliferation disorder related sequences and their complements; and as primers for the amplification of haematopoietic cell proliferation disorder related DNA sequences. The nucleotide sequences from the present invention can also be used for detecting a predisposition to, differentiation between subclasses, diagnosis, prognosis, treatment and/or monitoring of haematopoietic cell proliferative disorders. The present method enables a highly specific classification of haematopoietic cell proliferative

CC disorders allowing for improved and informed treatment of patients
 XX Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
 SQ Query Match 3.0%; Score 50.2; DB 7; Length 8056;
 Best Local Similarity 53.3%; Pred. No. 0.083;
 Matches 106; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 QY 35 TGAATAATTTTACCTAGTACTTCTGTAATCAATTTTTCAGAGGATGAAAAAATCG 94
 DB 1527 TAAAAAATTTTAAATAAATTAATTTTAAAAAATTAATTTTAAATTTTAAATTTTAA 1586
 QY 95 CAAGAAGCAAAATTTTAAATGAATGATGCAATATACAAATTTTAAATTTACACAAATTTG 154
 DB 1587 ATTATATTTTAAATTTTAAATTTTAAAAAATTTGAAATTAATGAAATTTGAAATTTAA 1646
 QY 155 AGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCATATTAATGCTAAAT 214
 DB 1647 AAAATTTTAAATTAATATGATAAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1706
 QY 215 AAAATTTCTATTTCTATTTT 233
 DB 1707 AATATAATTAATATATTT 1725
 RESULT 14
 ID AAS45498/c
 XX AAS45498 standard; DNA; 17142 BP.
 AC AAS45498;
 DT 18-DEC-2001 (first entry)
 DE Chemically pretreated genomic DNA associated with cell cycle #102.
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritis;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 OS Homo sapiens.
 PN WO200168911-A2.
 XX 20-SEP-2001.
 XX 15-MAR-2001; 2001WO-EP002945.
 XX 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2001-602751/68.
 DR Designing primers and probes for analyzing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle.
 XX Claim 1; SEQ ID NO 203; 28pp; English.
 CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing

CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers
 SQ Sequence 17142 BP; 5054 A; 299 C; 3741 G; 8048 T; 0 U; 0 Other;
 Query Match 3.0%; Score 50.2; DB 4; Length 17142;
 Best Local Similarity 50.6%; Pred. No. 0.11;
 Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
 QY 17 ATACCACTCATCTACATGTCATAAATTTTATAGTTAGTACTTCTGTAATCAATTTAG 76
 DB 9510 ATATACAAATATCTATAAATTAACCTTTAAAAAACAACAACTTTTATTCAAAAATCTCCA 9451
 QY 77 AGGATGAAAAAATTCGCAAGAACAAATATTTTAAATGAATGATGCAATATACAAAT 136
 DB 9450 CCTATAATTTAAAACTTAAAAAATACTAACTTAAATCTAACCTTCTAACAAA 9391
 QY 137 TTAATTACAAATTAATGTAAGATTACATTTAGTTTTCATAGAAATCAATTTCTAGAGT 196
 DB 9390 AAAAACTTTAAAAATTTAAATATCTACCACTAATTTATTTATTTATTTTAAATTT 9331
 QY 197 CATATAATGCTTAATTTAAATTTCTATTTCTATTTTCTTAAAGAAAAAAGAGCCCA 255
 DB 9330 CTTAATTTATATTAATTAATAAATTTTCCAAAAATAATTTAGTAAAAAATAATATCA 9272
 RESULT 15
 ID ABL34107/c
 XX ABL34107 standard; DNA; 17142 BP.
 AC ABL34107;
 DT 26-MAR-2002 (first entry)
 DE Human immune system associated gene SEQ ID NO: 2080.
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.
 OS Homo sapiens.
 PN WO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP007537.
 XX 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPIG-) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 DR Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:31:08 ; Search time 144 Seconds
(without alignments)
6389.643 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagggtcacataac.....ctcgtagttgggacggcg 1658

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	53.6	3.2	2750	3	US-08-617-860B-33	Sequence 33, Appl
C 2	51.4	3.1	6265	4	US-09-129-112-3	Sequence 3, Appl
C 3	48.8	2.9	8961	4	US-10-204-708-80	Sequence 80, Appl
C 4	46.4	2.8	2448	4	US-09-134-001C-777	Sequence 777, Appl
C 5	46.4	2.8	9347	4	US-10-204-708-35	Sequence 35, Appl
C 6	46	2.8	20674	4	US-09-641-638-651	Sequence 651, Appl
C 7	44.8	2.7	5852	1	US-07-867-106-2	Sequence 2, Appl
C 8	44	2.7	6656	4	US-10-204-708-75	Sequence 75, Appl
C 9	44	2.7	168575	4	US-09-426-290-1	Sequence 1, Appl
C 10	43.8	2.6	5152	4	US-10-204-708-74	Sequence 74, Appl
C 11	43.8	2.6	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C 12	43.4	2.6	6182	4	US-10-204-708-87	Sequence 87, Appl
C 13	43	2.6	1296	4	US-09-107-532A-3402	Sequence 3402, Ap
C 14	43	2.6	1296	4	US-09-134-000C-3111	Sequence 3111, Ap
C 15	43	2.6	5852	1	US-07-867-106-2	Sequence 2, Appl
C 16	43	2.6	10640	4	US-09-417-485D-5	Sequence 5, Appl
C 17	42.6	2.6	1850	3	US-08-617-860B-32	Sequence 32, Appl
C 18	42.6	2.6	4098	2	US-08-605-106-4	Sequence 4, Appl
C 19	42.6	2.6	9636	1	US-08-323-170B-1	Sequence 1, Appl
C 20	42.6	2.6	9636	4	US-08-954-441-1	Sequence 1, Appl
C 21	42.6	2.6	51952	3	US-08-947-823-1	Sequence 1, Appl
C 22	42.4	2.6	11049	4	US-10-204-708-22	Sequence 22, Appl
C 23	42.2	2.5	19025	4	US-09-849-334-3	Sequence 3, Appl
C 24	42.2	2.5	19025	4	US-10-274-878-3	Sequence 3, Appl
C 25	42	2.5	565	4	US-08-956-171E-1007	Sequence 3, Appl
C 26	42	2.5	6669	4	US-10-204-708-6	Sequence 6, Appl
C 27	41.8	2.5	11015	4	US-10-204-708-56	Sequence 56, Appl

C 28	41.8	2.5	53332	4	US-09-801-861-3	Sequence 3, Appl
C 29	41.6	2.5	1422	1	US-08-319-704-5	Sequence 5, Appl
C 30	41.6	2.5	5562	4	US-10-204-708-63	Sequence 63, Appl
C 31	41.4	2.5	6152	3	US-08-973-462-1	Sequence 1, Appl
C 32	41.4	2.5	640681	4	US-09-790-988-1	Sequence 1, Appl
C 33	41.4	2.5	1664976	4	US-08-916-421B-1	Sequence 1, Appl
C 34	41	2.5	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 35	40.8	2.5	672	4	US-09-134-001C-1992	Sequence 1992, Ap
C 36	40.8	2.5	6070	4	US-10-204-708-10	Sequence 10, Appl
C 37	40.8	2.5	580073	4	US-08-545-528D-1	Sequence 1, Appl
C 38	40.6	2.4	1784	4	US-09-601-198-21	Sequence 21, Appl
C 39	40.4	2.4	357	4	US-09-328-352-538	Sequence 538, App
C 40	40.4	2.4	1450	3	US-08-617-860B-22	Sequence 22, Appl
C 41	40.4	2.4	6768	1	US-08-107-755A-1	Sequence 1, Appl
C 42	40.4	2.4	8093	4	US-10-204-708-31	Sequence 31, Appl
C 43	40.4	2.4	8457	1	US-07-991-867B-1	Sequence 1, Appl
C 44	40.4	2.4	8457	2	US-08-544-332-1	Sequence 1, Appl
C 45	40.4	2.4	8457	4	US-09-370-861A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-617-860B-33/c
; Sequence 33, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filsak, E.,
; APPLICANT: Hvricke-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Mller, A., Schulte, W., Voetz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
; STREET: 1140 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,860B
; FILING DATE: 01-VAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02950
; FILING DATE: 05-SEP-1994
; APPLICATION NUMBER: DE P4329951.2
; FILING DATE: 04-SEP-1993
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2750 Base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: CITE94
; FEATURE:
; NAME/KEY: Startcodon
; LOCATION: 2637..2639
; FEATURE:
; NAME/KEY: CDS

[illegible]

Sequence 777, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 777
LENGTH: 2448
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-777

Query Match 2.8%; Score 46.4; DB 4; Length 2448;
Best Local Similarity 52.0%; Pred. No. 0.017;
Matches 104; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 26 ATCTACATGTGATAAATTTTATAGTTAGATTACTTCTGTAATCATTTTCAGAGGATGAA 85
Db ATCTACCTCGTACATTTGTATGAAACACACAAAATTTCTCAATGAGATATTGAGA 802

QY 86 AAAAAATCGAAGAAGCAAAATATTTTAAATGAATGATGCAATATACAAAATTTAATTA 145
Db AACCAATTGAAACAGAGCAATTTACTAGAAATCATTTGACGTTTATCAGAGAGAAAAAT 862

QY 146 CAATATGATGATACATTTTCTTTAGTTTCTATGATGATGATGATGATGATGATGATGAT 205
Db TAATATCTGAAGCTGATCAGGTATCAATTTCCAAAGTTTATCTACTATTTCAGAAATGAAAAGTG 922

QY 206 GCCTAAATTTAAATTTCTATT 225
Db TGCAAACTTATACCAATTT 942

RESULT 5
US-10-204-708-35/c
Sequence 35, Application US/10204708
Patent No. 667731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 35
LENGTH: 9347
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35

Query Match 2.8%; Score 46.4; DB 4; Length 9347;
Best Local Similarity 54.8%; Pred. No. 0.035;
Matches 92; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 83 AAAAAAATCGAAGAAGCAAAATATTTTAAATGAATGATGCAATATACAAAATTTAAT 142
Db AAAAAAATACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7627

QY 143 ACACAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 202
Db ACACATTTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 7567

QY 203 AATGCTTAAATTTAATTTCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 250
Db AAAAAAATACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7519

RESULT 6
US-09-641-638-651/c
Sequence 651, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GENSET.051CP1
CURRENT APPLICATION NUMBER: US/09/641,638
CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 651
LENGTH: 20674
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1123..3123
OTHER INFORMATION: 5' regulatory region
NAME/KEY: exon
LOCATION: 3124..3297
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 3871..4072
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 5552..5633
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 5758..5880
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 5996..6099
OTHER INFORMATION: exon 5
NAME/KEY: exon
LOCATION: 6349..6509
OTHER INFORMATION: exon 6
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8

```
, NAME/KEY: exon
, LOCATION: 12254..12340
, OTHER INFORMATION: exon 9
, NAME/KEY: exon
, LOCATION: 12854..13023
, OTHER INFORMATION: exon 10
, NAME/KEY: exon
, LOCATION: 13308..13429
, OTHER INFORMATION: exon 11
, NAME/KEY: exon
, LOCATION: 16567..16667
, OTHER INFORMATION: exon 12
, NAME/KEY: exon
, LOCATION: 16775..16945
, OTHER INFORMATION: exon 13
, NAME/KEY: exon
, LOCATION: 17063..17554
, OTHER INFORMATION: exon 14
, NAME/KEY: misc feature
, LOCATION: 17555..20674
, OTHER INFORMATION: 3'regulatory region
, NAME/KEY: allele
, LOCATION: 1128
, OTHER INFORMATION: 10-508-191 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 1182
, OTHER INFORMATION: 10-508-245 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 1559
, OTHER INFORMATION: 10-509-284 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 1570
, OTHER INFORMATION: 10-509-295 : deletion of C
, NAME/KEY: allele
, LOCATION: 1827
, OTHER INFORMATION: 10-510-173 : variable motif ATTTA or TTTTTT
, NAME/KEY: allele
, LOCATION: 2048
, OTHER INFORMATION: 10-511-62 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 2323
, OTHER INFORMATION: 10-511-337 : insertion of T
, NAME/KEY: allele
, LOCATION: 2341
, OTHER INFORMATION: 10-512-36 : polymorphic base G or C
, NAME/KEY: allele
, LOCATION: 2623
, OTHER INFORMATION: 10-512-318 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 2832
, OTHER INFORMATION: 10-513-250 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 2844
, OTHER INFORMATION: 10-513-262 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 2934
, OTHER INFORMATION: 10-513-352 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 2947
, OTHER INFORMATION: 10-513-365 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 3802
, OTHER INFORMATION: 12-206-81 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 4062
, OTHER INFORMATION: 10-343-231 : deletion of C
, NAME/KEY: allele
, LOCATION: 4088
, OTHER INFORMATION: 12-206-366 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 4109
, OTHER INFORMATION: 10-343-278 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 4170
, OTHER INFORMATION: 10-343-339 : polymorphic base G or T
, NAME/KEY: allele
, LOCATION: 5903
, OTHER INFORMATION: 10-346-23 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 6019
, OTHER INFORMATION: 10-346-141 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 6141
, OTHER INFORMATION: 10-346-263 : polymorphic base G or C
, NAME/KEY: allele
, LOCATION: 6183
, OTHER INFORMATION: 10-346-305 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 6338
, OTHER INFORMATION: 10-347-74 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 6375
, OTHER INFORMATION: 10-347-111 : polymorphic base G or C
, NAME/KEY: allele
, LOCATION: 6429
, OTHER INFORMATION: 10-347-165 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 6467
, OTHER INFORMATION: 10-347-203 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 6484
, OTHER INFORMATION: 10-347-220 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 6534
, OTHER INFORMATION: 10-347-271 : polymorphic base A or T
, NAME/KEY: allele
, LOCATION: 6611
, OTHER INFORMATION: 10-347-348 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 7668
, OTHER INFORMATION: 10-348-391 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 8608
, OTHER INFORMATION: 10-349-47 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 8658
, OTHER INFORMATION: 10-349-97 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 8703
, OTHER INFORMATION: 10-349-142 : polymorphic base G or C
, NAME/KEY: allele
, LOCATION: 8777
, OTHER INFORMATION: 10-349-216 : deletion of CTG
, NAME/KEY: allele
, LOCATION: 8785
, OTHER INFORMATION: 10-349-224 : polymorphic base G or T
, NAME/KEY: allele
, LOCATION: 8926
, OTHER INFORMATION: 10-349-368 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 12171
, OTHER INFORMATION: 10-350-72 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 12429
, OTHER INFORMATION: 10-350-332 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 13341
, OTHER INFORMATION: 10-507-170 : polymorphic base A or G
, NAME/KEY: allele
, LOCATION: 13492
, OTHER INFORMATION: 10-507-321 : polymorphic base A or C
, NAME/KEY: allele
, LOCATION: 13524
, OTHER INFORMATION: 10-507-353 : polymorphic base C or T
, NAME/KEY: allele
, LOCATION: 13535
```


Db 1138 AATATAATTAAATCTATTACTCACAAATCTTTCAATATACTAAAT 1091

RESULT 9

US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1

; LENGTH: 168575

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (21181)...(21403)

; NAME/KEY: CDS

; LOCATION: (95252)...(95430)

; NAME/KEY: CDS

; LOCATION: (101753)...(101996)

; NAME/KEY: CDS

; LOCATION: (110324)...(110439)

; NAME/KEY: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

US-09-426-290-1

Query Match 2.7%; Score 44; DB 4; Length 168575;

Best Local Similarity 47.8%; Pred. No. 0.74; Indels 0; Gaps 0;

Matches 128; Conservative 0; Mismatches 140; Mismatches 140; Indels 0; Gaps 0;

QY 29 TACATGTGATAAATTTTAGTTAGTACTCTTGTAAATCAATTCAGAGGATGAAAAA 88

Db 109154 TATATATATATATAATATAATAAGTTATATATATAATAATAATAAGTTATATAT 109095

QY 89 AATCCAGAAAGCAAAATTTTAAATGAATGATGCAATATACAAATTTAATTACAA 148

Db 109094 ATATATTTAGATATAATAAATAATATATATATTAATAGATATTTATATGTAATATAT 109035

QY 149 TTATGTAAGATTACATTTGTTTGTTCATAGAAATCAATTTCTAGAGTCATAAATGCC 208

Db 109034 ATAACTATATATACATTAATATTTATATATATATATATATATATATATATATATAT 108975

QY 209 TAAATTAATCTATTCTATTTTCTTAAAGAAAAAAGCCCATTAAGGACCAT 268

Db 108974 TATTTATATATAATATATATAAATATAATATATATATATATATATATATATATATAT 108915

QY 269 AGAATGGCCCTGCTCAATTTTATATA 296

Db 108914 TGAATTTTGTATGTATTTTATTTT 108887

RESULT 10

US-10-204-708-74/c

; Sequence 74, Application US/10204708

; Patent No. 6677731

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

; FILE REFERENCE: 5013.1012

; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 74
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-74

Query Match 2.6%; Score 43.8; DB 4; Length 5152;

Best Local Similarity 51.8%; Pred. No. 0.13; Indels 0; Gaps 0;

Matches 99; Conservative 0; Mismatches 92; Mismatches 92; Indels 0; Gaps 0;

QY 83 AAAAAAATCGCAAGAAACAAATATTTTAAATGATGATGCAATATACAAATTTAAAT 142

Db 4914 AAATTAATTTCTTTTAAACAAACATATATATATATATATATATATATATATATAT 4855

QY 143 ACACAAATATGTAAGATTACATGTTTGTTCATAGAAATCAATTTCTAGAGTCATAT 202

Db 4854 AAATAACATATATTTTAAACAATAAATACCTTATTTTAAACCTTATTTTATTTTTC 4795

QY 203 AATGCTAAATTAATTTCTATTTTCTTAAAGAAAAAAGCCCATTAAGGG 262

Db 4794 TATACTTAACATAACCTCATTTATAATATATACCTAAAAAATTAATATCCATAAAC 4735

QY 263 ACCATTAGAAT 273

Db 4734 AAATTAAT 4724

RESULT 11

US-08-916-421B-1

; Sequence 1, Application US/08916421B

; Patent No. 6503729

; GENERAL INFORMATION:

; APPLICANT: Bult et al.

; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus

; Patent No. 6503729

; FILE REFERENCE: jannaschii

; CURRENT APPLICATION NUMBER: US/08/916,421B

; CURRENT FILING DATE: 1997-08-22

; PRIOR APPLICATION NUMBER: US 60/024,428

; PRIOR FILING DATE: 1996-08-22

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1664976

; TYPE: DNA

; ORGANISM: Methanococcus jannaschii

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (28222)..(28222)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (28257)..(28258)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

; LOCATION: (84773)..(84773)

; OTHER INFORMATION: n equals a, t, c, or g

; NAME/KEY: misc feature

```

; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g

```


[illegible]

```

RESULT 14
US-09-134-000C-3111
; Sequence 3111, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3111
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3111

```

Query Match	2.6%;	Score 43;	DB 4;	Length 1296;
Best Local Similarity	49.3%;	Pred. No. 0.099;		
Matches 112;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0
QY	41	ATTTTATAGTTAGTATTACTTCTTCTTAATCATTTTCAGAGGATGAAAAAAAATTCGACAA	100	
Db	408	ATTTAGTAGTGGAGTATTTCACATTAAGAAGGTAAATAATATAGAAGAAATATGATAAGA	467	
QY	101	AGCAATATTTTTAAATCAATGATGCAATATACAAATTTTAAATTCACAAATTTATGTAAAGATT	160	
Db	468	TTCAATATATTGTTTCATGAAGAATTTGCTAAACAAACAACTCTAAAAATTAGGTGATGAAGT	527	
QY	161	ACATTGTTTAGTTTCATAGAAAACCAATTTCTAGAGTCATATATATGCTTAATTTAAAAATT	220	
Db	528	TGATCTTGAATTTACTGATATTGAAAAAAGTGGAAAAATAAAAAAGTCAATAAATTTAAAAAT	587	
QY	221	CTATCTTATTTTTTCTTTAAGAAAAAAAACACGCCCATTTAAGGGACCAT	267	
Db	588	TATAGGATCTTTTCTCGTAAAAAACACAGGAACATATACAGGATTTAT	630	

```

RESULT 15
US-07-869-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
;
; GENERAL INFORMATION:
;
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
;
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; Slime Moulds of the Genus Dictyostelium
;
; NUMBER OF SEQUENCES: 19
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526/ris
;
; STREET: One liberty Place 46th Floor
;
; CITY: Philadelphia
;
; STATE: PA
;
; COUNTRY: USA
;
; ZIP: 19103
;
; COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

```

	Query Match	2.8%	Score 43;	DB 1;	Length 5852;
	Best Local Similarity	54.8%;	Pred. No. 0.22;		
	Matches	85; Conservative	0; Mismatches	70; Indels	0; Gaps
QY	83	AAAAAAAAATCGCAAGAAAGCAAAATATTTTAAATGATGATGCAATATATACAAATTTAAAT	142		
Db	1850	AAAGAAAAAATAAAAAAGTAGAATTTATTAATAAAATTTAAATATTTATTCATCTCTTAATA	1791		
QY	143	ACACAATTATGTAAAGATTACATTTGTTTAGTTTCATAGAAATCAATTTCTAGAGTCATAAT	202		
Db	1790	AATTAGTATATATCATAGGCAATTTATTTTATCTATCTCTAAAAAAAACCTAGGAAA	1731		
QY	203	AATGCCTAAATTTAAATCTATCTATTTTTCCTT	237		
Db	1730	AATGAATGTCATCAAAATAGTATTTTAAACATTTT	1696		

Search completed: September 2, 2004, 10:46:19
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 09:38:09 ; Search time 785 Seconds
(without alignments)
10498.497 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcaggctcaataac.....ctcgttagttggagcggc 1658

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	17	US-10-732-721-1
2	98.4	5.9	836	13	US-10-425-114-7252
3	97.4	5.9	918	13	US-10-425-114-13779
4	71.4	4.3	894	13	US-10-425-114-3295
5	53.2	3.2	7498	15	US-10-311-455-230
6	52.4	3.1	6012	13	US-10-221-613-236
7	51.4	3.1	6265	9	US-09-129-112-3
8	50.8	3.1	12763	15	US-10-311-455-276
9	50.8	3.1	113515	15	US-10-311-455-2147
10	50.6	3.1	17280	13	US-10-221-714A-498
11	50.4	3.0	1134	15	US-10-074-475-28
12	50.4	3.0	15767	15	US-10-311-455-1179
13	50.4	3.0	15767	15	US-10-240-485-105
14	50.2	3.0	17142	15	US-10-239-676-206

C 15	50.2	3.0	17142	15	US-10-311-455-2080	Sequence 2080, Ap
C 16	50.2	3.0	17142	15	US-10-240-453-304	Sequence 304, App
C 17	49.6	3.0	653	13	US-10-027-632-104828	Sequence 104828, Ap
C 18	49.6	3.0	653	16	US-10-027-632-104828	Sequence 104828, Ap
C 19	49.6	3.0	778	13	US-10-027-632-142028	Sequence 142028, Ap
C 20	49.6	3.0	778	16	US-10-027-632-142028	Sequence 142028, Ap
C 21	49.6	3.0	3673778	15	US-10-312-841-1	Sequence 1, Appli
C 22	49.4	3.0	591	17	US-10-437-963-21828	Sequence 21828, A
C 23	49.4	3.0	14551	15	US-10-240-485-137	Sequence 137, App
C 24	48.8	2.9	8961	15	US-10-204-708-80	Sequence 80, Appl
C 25	48.8	2.9	8961	15	US-10-240-453-302	Sequence 302, App
C 26	48.6	2.9	5891	13	US-10-221-613-43	Sequence 43, Appl
C 27	48.4	2.9	632	9	US-09-772-134B-57	Sequence 57, Appl
C 28	48.4	2.9	736	9	US-09-772-134B-53	Sequence 53, Appl
C 29	48.2	2.9	5313	15	US-10-311-455-736	Sequence 736, App
C 30	48.2	2.9	5768	15	US-10-311-455-2136	Sequence 2136, Ap
C 31	48.2	2.9	7276	15	US-10-311-455-875	Sequence 875, App
C 32	48.2	2.9	13038	15	US-10-311-455-1248	Sequence 1248, Ap
C 33	48.2	2.9	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 34	48	2.9	6053	15	US-10-239-676-76	Sequence 76, Appl
C 35	48	2.9	6053	15	US-10-240-453-82	Sequence 82, Appl
C 36	48	2.9	7657	15	US-10-239-676-185	Sequence 185, App
C 37	48	2.9	7657	15	US-10-311-455-1995	Sequence 1995, Ap
C 38	48	2.9	33053	17	US-10-433-793-36	Sequence 36, Appl
C 39	47.8	2.9	3673778	15	US-10-312-841-1	Sequence 1, Appli
C 40	47.6	2.9	5499	17	US-10-433-793-2	Sequence 2, Appli
C 41	47.6	2.9	9963	15	US-10-311-455-667	Sequence 667, App
C 42	47.6	2.9	18997	15	US-10-172-086-17	Sequence 17, Appl
C 43	47.6	2.9	18997	15	US-10-311-455-543	Sequence 543, App
C 44	47.6	2.9	18997	17	US-10-311-507-31	Sequence 31, Appl
C 45	47.6	2.9	113515	15	US-10-311-455-2148	Sequence 2148, Ap

ALIGNMENTS

RESULT 1

US-10-732-721-1
; Sequence 1, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-15(52826A)
; CURRENT APPLICATION NUMBER: US/10/732.721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match	100.0%;	Score 1658;	DB 17;	Length 1658;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1658;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GAATTCAGGCTCAATACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60	
Db	1	GAATTCAGGCTCAATACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60	
Qy	61	CTTCTAATCATTTTCAGAGGATGAAAAAATCCCAAGAAAGCAATATTTTAATGAAT	120	
Db	61	CTTCTAATCATTTTCAGAGGATGAAAAAATCCCAAGAAAGCAATATTTTAATGAAT	120	
Qy	121	GATCCAATATACAAATTTAATTAATGACAAATTTATGTAAGATTACATTTGTTTTCATAGA	180	
Db	121	GATCCAATATACAAATTTAATTAATGACAAATTTATGTAAGATTACATTTGTTTTCATAGA	180	

QY 181 AATCAATTTCTAGATGCTAATAATGCTTAAATTTAAATTTCTATTTCTTTTCTTAAAG 240
 DB 181 AATCAATTTCTAGATGCTAATAATGCTTAAATTTAAATTTCTATTTCTTTTCTTAAAG 240
 QY 241 AAAAAAAGAGCCCAATTAAGGACCAATTAAGATGCGGCTGCTCCATTTTATATATA 300
 DB 241 AAAAAAAGAGCCCAATTAAGGACCAATTAAGATGCGGCTGCTCCATTTTATATATA 300
 QY 301 GAGATATAGTTGTTGCTCACTAGGCCACTCCACCTCGCATGCGTCTTTATATATAG 360
 DB 301 GAGATATAGTTGTTGCTCACTAGGCCACTCCACCTCGCATGCGTCTTTATATATAG 360
 QY 361 AAAAAATGAGCGGCTGCTTTGCAATTAAGGCTTACAGATGACGAGTCAAGCTTCGT 420
 DB 361 AAAAAATGAGCGGCTGCTTTGCAATTAAGGCTTACAGATGACGAGTCAAGCTTCGT 420
 QY 421 ATTCACTGACGAGTAATTTTAAATGACGAGGCGCTTTGTAATTCCTCAGCAGCT 480
 DB 421 ATTCACTGACGAGTAATTTTAAATGACGAGGCGCTTTGTAATTCCTCAGCAGCT 480
 QY 481 CTGGGCGCGGCTGCTGCTGCTAGCGGAGTGAAGAGCAGCAGCAGCTCTACGTGCCACG 540
 DB 481 CTGGGCGCGGCTGCTGCTGCTAGCGGAGTGAAGAGCAGCAGCAGCTCTACGTGCCACG 540
 QY 541 TGCCTGCGCTGCTGCGCCACCGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 541 TGCCTGCGCTGCTGCGCCACCGTGAAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 AGCGACGAGTTGAAGCTAGCGGCTGGCTGCTGCTGCGCGCTGCTAGCACGCGGAC 660
 DB 601 AGCGACGAGTTGAAGCTAGCGGCTGGCTGCTGCTGCGCGCTGCTAGCACGCGGAC 660
 QY 661 GAGCTGAGCTGATGCTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB 661 GAGCTGAGCTGATGCTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 721 ATGCACTAAATAAAGTGTATCTCATGTGATGATGATGATGATGATGATGATGATGATG 780
 DB 721 ATGCACTAAATAAAGTGTATCTCATGTGATGATGATGATGATGATGATGATGATGATG 780
 QY 781 ATCAGTAACAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 781 ATCAGTAACAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 841 ATTTGCTGCTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 DB 841 ATTTGCTGCTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 901 GAGTTCAACATCTTTCTTAATACCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 GAGTTCAACATCTTTCTTAATACCTTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 TAACGTTTAAAGACCCCTTTGTAACGAGTGTAGTCCAGGCTGCTTAACGTAATAATAG 1020
 DB 961 TAACGTTTAAAGACCCCTTTGTAACGAGTGTAGTCCAGGCTGCTTAACGTAATAATAG 1020
 QY 1021 TAGAACACAGGAGACAGAGTGTAGAGGAACTGATTTCTTTGTTACTATATAGTGGCTG 1080
 DB 1021 TAGAACACAGGAGACAGAGTGTAGAGGAACTGATTTCTTTGTTACTATATAGTGGCTG 1080
 QY 1081 CTCTCAAAGGTTTACATGATATGGGATCTCTCTCTATTTATAGACAACTAGGTTT 1140
 DB 1081 CTCTCAAAGGTTTACATGATATGGGATCTCTCTCTATTTATAGACAACTAGGTTT 1140
 QY 1141 CAGGCTATGGGCGACATAGGCTTCTGCGCCCAAGAAAGGTTTCTTAACACTACCATCT 1200
 DB 1141 CAGGCTATGGGCGACATAGGCTTCTGCGCCCAAGAAAGGTTTCTTAACACTACCATCT 1200
 QY 1201 TCCAGCGCGGCTGCTGCTGCTTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 DB 1201 TCCAGCGCGGCTGCTGCTGCTTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 QY 1261 GGGTCTACGCTGACGCGCACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

DB 1261 GGGTCTACGCTGACGCGCACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 QY 1321 CGAGGCGCATCTGCTGACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 DB 1321 CGAGGCGCATCTGCTGACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
 QY 1381 GTGCGCGCGACGCGCGCGGAGGACCTTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 DB 1381 GTGCGCGCGACGCGCGCGGAGGACCTTGCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 QY 1441 GGTCCGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 DB 1441 GGTCCGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 QY 1501 CGGCTATAAAGAGGCGCGCCACCTGATCTCTCATCTCAACAAGAAAGCAGCAGCAAC 1560
 DB 1501 CGGCTATAAAGAGGCGCGCCACCTGATCTCTCATCTCAACAAGAAAGCAGCAGCAAC 1560
 QY 1561 AGCCACACCTAACTAAAGAACAGTAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 DB 1561 AGCCACACCTAACTAAAGAACAGTAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 QY 1621 CCGAGCTCGCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
 DB 1621 CCGAGCTCGCCAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658

RESULT 2

US-10-425-114-7252
 ; Sequence 7252, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 7252
 ; LENGTH: 836
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700623908_FLI
 US-10-425-114-7252

Query Match 5.9%; Score 98.4; DB 13; Length 836;
 Best Local Similarity 99.0%; Pred. No. 7.4e-14;
 Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACGCCAACACCTAAACTAAAGAACAGTAGTCCCTGCTAGCTAGCAAC 1612
 DB 1 GCAGCAACGCCAACACCTAAACTAAAGAACAGTAGTCCCTGCTAGCTAGCAAC 60
 QY 1613 GATCCGCGCGAGCTCGCCAGTTTGTCTGCTAGTTTGGG 1652
 DB 61 GATCCGCGCGAGCTCGCCAGTTTGTCTGCTAGTTTGGG 100

RESULT 3

US-10-425-114-13779
 ; Sequence 13779, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua


```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13779
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-042-D8_FLI
US-10-425-114-13779

Query Match          5.9%; Score 97.4; DB 13; Length 918;
Best Local Similarity 99.0%; Pred. No. 1.4e-13;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1554 CAGCAACACCCACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAACG 1613
DB 1 CAGCAACACCCACACCTAACTAAAGAACAGTAGTCCCTGTGTAGCTCTAGCAACG 60
QY 1614 ATCCGCGCGAGCTGCCAGTTTGTCTCGTGTAGTTGGG 1652
DB 61 ATCCGCGCGAGCTGCCAGTTTGTCTCGTGTAGTTGGG 99

RESULT 4
US-10-425-114-3295
; Sequence 3295, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3295
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258323_FLI
US-10-425-114-3295

Query Match          4.3%; Score 71.4; DB 13; Length 894;
Best Local Similarity 98.6%; Pred. No. 4.4e-07;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTGTCTGTGTAGCTAGCAACGATCCGCGCGAGCTCGCCAGTTTGGC 1639
DB 1 GAACAGTAGTGTCTGTGTAGCTAGCAACGATCCGCGCGAGCTCGCCAGTTTGGC 60
QY 1640 TCGTTAGTTGGG 1652
DB 61 TCGTTAGTTGGG 73

RESULT 5
US-10-311-455-230/c
; Sequence 230, Application US/10311455
; Publication No. US20030143606A1

```

```

; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 230
; LENGTH: 7498
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-230

Query Match          3.2%; Score 53.2; DB 15; Length 7498;
Best Local Similarity 50.4%; Pred. No. 0.066;
Matches 130; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 12 TCACAATACCGTCATCTACATGTGATAAATTTTATAGTTAGATTACTTCTTGTATCAT 71
DB 1100 TCTAAATAATAATTATATATATATATCTTAAAAATAAATAATATATATATATTTTCTTAAAA 1041
QY 72 TTCAGAGGATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATA 131
DB 1040 TAAAAATATATATATATAATCTAAAAATAACAATATATATATATATATATCTATAATAA 981
QY 132 CAAATTTAATTACAAATATGTAGATACATTTGTTAGTTTCATAGAAATCAATTTCT 191
DB 980 AATATATATATAATCTAAAAATAAAAAATATATATATATAATCTAAAAATAAAAAATATAT 921
QY 192 AGAGTCATAATAATCGCTAAATTTAAATTTCTATCTTCTTTTCTTAAAGAAAAACAG 251
DB 920 ATAAATATAAAAAATAAAAAATAAAAAATATCCCTCTAAAAATAAAAAATAAAAAATAA 861
QY 252 CCATTAAAGGGACCATTA 269
DB 860 CTAAATAACAACCCATTA 843

RESULT 6
US-10-221-613-236/c
; Sequence 236, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01

```

NUMBER OF SEQ ID NOS: 428
SEQ ID NO 236
LENGTH: 6012
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-236

Query Match 3.1%; Score 52.2; DB 13; Length 6012;
Best Local Similarity 50.2%; Pred. No. 0.1;
Matches 129; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 3 ATTACGGCTCAACATACAGTCACTACATGATGATAAATTTATAGTTAGATTACTTCT 62
DB 1299 ATTCAATTTCAAATTTACCAAAATATACACCAATTTTATATAAATTTCTCTATTAT 1240

QY 63 TGTAAATCATTTCAGAGATGCAAAAAAATCGCAAGAAAGCAATATTTTAAATGAATGA 122
DB 1239 TCATTCCCTTATCATTCACAAAAATTAACAACAAATTAACAACTATTAATTTCAATA 1180

QY 123 TGCATATACAAATTTAATTACAAATATATGTAAGATTACATTTGTTTTCATAGAAA 182
DB 1179 CCTACAAATAAATTAACAATAACCACTCCATAAAAAATTTCTATTACGATAAAAAAT 1120

QY 183 TCAATTTCTAGAGTCATAATAATGCCCTAAATTAATTTCTATTCTTTTCTTAAAGAA 242
DB 1119 CCAATAAATACACGATTTTCTTCTTATTTTCTTTTATTCATCTCTTTTACATTCATA 1060

QY 243 AAAAAACAGCCCATTA 259
DB 1059 ATATTCCCGCTCTAA 1043

RESULT 7
US-09-129-112-3/c
Sequence 3, Application US/09129112
Patent No. US2002001995A1
GENERAL INFORMATION:
APPLICANT: Etxler, Marilyn B.
APPLICANT: Murphy, Judith B.
TITLE OF INVENTION: The Regents of the University of California
FILE REFERENCE: 023070-079810US
CURRENT APPLICATION NUMBER: US/09/129,112
PRIOR FILING DATE: 1998-08-04
PRIOR FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 6265
TYPE: DNA
ORGANISM: Dolichos biflorus
FEATURE:
OTHER INFORMATION: genomic sequence of NBP46 (DB46)

NAME/KEY: exon
LOCATION: (633)..(944)
NAME/KEY: intron
LOCATION: (945)..(1022)
NAME/KEY: exon
LOCATION: (1023)..(1151)
NAME/KEY: intron
LOCATION: (1152)..(1559)
NAME/KEY: exon
LOCATION: (1560)..(1616)
NAME/KEY: intron
LOCATION: (1617)..(1697)
NAME/KEY: exon
LOCATION: (1698)..(1790)
US-09-129-112-3

Query Match 3.1%; Score 51.4; DB 9; Length 6265;

Best Local Similarity 49.4%; Pred. No. 0.16;
Matches 133; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 24 TCATCTACATGTCATAAATTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGATGA 83
DB 3725 TCCGCTTAAAGGTAAAAATTTGCAGAAAAAATACATATGTATTAATTAATATATAT 3666

QY 84 AAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATACAAAAATTAATTA 143
DB 3665 AATATAACATATAAATATTAATATTTTATTTATTTATTTATTTATTAATTAATAAATTA 3606

QY 144 CACAATTTATGTAAGATTACATTTGTTTGTTCATAGAAATCAATTTCTAGAGTCATAA 203
DB 3605 AATAATAATTAATTAATAATTTATTTATTTTATAATAATAAAATTTAAAAAATAATAA 3546

QY 204 ATGCCCTAAATTAATTTCTTCTATTCTTTTCCCTTAGAAAAAAGACGCCCATTAAGGGA 263
DB 3545 AATAATAAATTTAAAAATAAATTTTATTTTTCATAATAATAAATAAATAATTAATTT 3486

QY 264 CCATTAGAATGCGCGCTGCTCCATTTT 292
DB 3485 AATAATAAATAAATTAATTTCACTTTT 3457

RESULT 8
US-10-311-455-276/c
Sequence 276, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Cytosine
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 276
LENGTH: 12763
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-276

Query Match 3.1%; Score 50.8; DB 15; Length 12763;
Best Local Similarity 56.6%; Pred. No. 0.36;
Matches 94; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 83 AAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATACAAAAATTAAT 142
DB 2935 AAAAAAATTTAAATTAACCAACCATTAATATAATTAATAAATAAATAATTTTATT 2876

QY 143 ACACAATTTATGTAAGATTACATTTGTTTGTTCATAGAAATCAATTTCTAGAGTCATAA 202
DB 2875 TATTTATTAATTTTAATAATACCTTTTAAACAATTAATAAATTTCTTTCTATTCT 2816

QY 203 AATGCCCTAAATTAATTTCTTCTATTCTTTTCCCTTAAGAAAAA 248
DB 2815 AATTATAAATAAATAAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTT 2770

RESULT 9
US-10-311-455-2147/c
Sequence 2147, Application US/10311455
Publication No. US20030143606A1

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2147
; LENGTH: 113515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2147

Query Match          3.1%; Score 50.8; DB 15; Length 113515;
Best Local Similarity 51.8%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 115; Conservative 0; Mismatches 107;

QY 25 CATCTACATGTGATAAAATTTTATAGTTAGATTACTCTTGTAAATCAATTCAGAGGATGAA 84
DB 11462 CAAATATATATACAAATATATATATATACAAATATATATATTTAATATCTACATATA 11403

QY 85 AAAAAAATCGAAGAACCAATATTTTAAATGAATGATGCAATATACAAATTTAATTAC 144
DB 11402 AAAATAACCAAAATTAATAATATATTTTCTATTATTTTAAATATCTTATACAAATCAC 11343

QY 145 ACAATTATGTAGATTACATTTGTTAGTTTTCATAGAAATCAATTTCTAGAGTCATAATA 204
DB 11342 AATAAATAAAATTTTAATCATATTCAAAAAATAACCATATTTAAACACAAAAAAT 11283

QY 205 TGCCTAAATTAATTTCTATTCTTTTTCCTTAAGAAAAAA 246
DB 11282 TACAATAAATAATAAAATTTCTTTAAATCAAAATTCAAA 11241

RESULT 10
US-10-221-714A-498/c
; Sequence 498, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Tumor Suppressor Genes and Oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 498
; LENGTH: 17280
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-498

Query Match          3.1%; Score 50.6; DB 13; Length 17280;
Best Local Similarity 52.1%; Pred. No. 0.52;
Matches 113; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 38 TAAATTTTATAGTTAGATTACTCTTGTAAATCAATTCAGAGGATGAAAAAATTCGCAA 97
DB 5424 TAAATAACAAATATAAATTTTCATCTCAAAAAAATAAATAAATAAATAAATAAATAA 5365

QY 98 GAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTTAAATTTACAAATTTATGTAA 157
DB 5364 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5305

QY 158 ATTACATTTGTTAGTTTCATAGAATCAATTTCTAGAGTCATATAATGCCTAAATTTAA 217
DB 5304 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5245

QY 218 ATTCTATTCTATTTTTCCTTAAGAAAAAATTAAGAAAAAATTAAGAAAAAATTAAGAAAA 254
DB 5244 CAACATACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5208

RESULT 11
US-10-074-475-28
; Sequence 28, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-28

Query Match          3.0%; Score 50.4; DB 15; Length 1134;
Best Local Similarity 51.3%; Pred. No. 0.091;
Matches 117; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 11 CTCACAATACCAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTCTTGTATCA 70
DB 29 CTGAAAGTTCTAGTCAATTTTAATTTGATCCAAATAGTTTTCTGAAATCTCTTTTAAGT 88

QY 71 TTTCCAGAGGATGAAAAAATAATTCGAAAGCAAAATATTTTAAATGAATGATCAATAT 130
DB 89 TCCAAGAAATTTCTATTATAAATAAGTGTACTTTTACCAATTTCCATTGTATAGCAACAG 148

QY 131 ACAAAATTAATTAACAATTAATGAATGATGATGATGATGATGATGATGATGATGATGATG 190
DB 149 ACACCTTTTAGAAAAGGATAAGTAATCAATCAATTTGTTTTTTTTTAAAAAATAAATTC 208

QY 191 TAGAGTCATATAATGCCTAAATTAATTAATTTCTATTTCTTTCTTA 238
```

Db 209 CAGACTACTAAATTTGGCATAAGAATAATTCITTTAAATGCAACATA 256

RESULT 12

US-10-311-455-1179
; Sequence 1179, Application US/10311455
; Publication No. US20030143606A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311,455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 1179

; LENGTH: 15767

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1179

Query Match 3.0%; Score 50.4; DB 15; Length 15767;

Best Local Similarity 49.3%; Pred. No. 0.55;

Matches 132; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 33 TGTGATAAAATTTATAGTTAGTACTTCTTGTAAATCAATTCAGAGGATGAAAAAAAT 92

Db 14988 TATTTTGAGATGGAAGTTAAATTTAAATTTTGAAGTTAAAGTTAAATTTATTTT 15047

QY 93 CGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATACAAATAT 152

Db 15048 TGTGAAGAATATTTTATAGATAATGTTGTATATTTATATTTTATATACGAAAGTAT 15107

QY 153 GTAAGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCAATAATGCGCTAAA 212

Db 15108 GTAAATGTTGATTTGTTTATTTTGGTGAATGTTTATTAAGTCAATTTGTTGTTTAAA 15167

QY 213 TTAATAATTCATTTCTATTTTCCCTTAAGAAAAAAACAGCCCATTAAGGACCATTAGAA 272

Db 15168 TTGAATTTAAATATTTTATTTTAAATTTAAATTTGTAATAATTAAGTGTATTAGAG 15227

QY 273 TGGCGCTGCTCCATTTTATATATA 300

Db 15228 AAGTTTATATTTATTTTGTGTTTTTA 15255

RESULT 13

US-10-240-485-105
; Sequence 105, Application US/10240485
; Publication No. US20030148327A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with

; FILE REFERENCE: 5013.1007

; CURRENT APPLICATION NUMBER: US/10/240,485

; CURRENT FILING DATE: 2002-10-02

; PRIOR APPLICATION NUMBER: PCT/EP01/03970

; PRIOR FILING DATE: 2001-04-06

; PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 202

; SEQ ID NO 105

; LENGTH: 15767

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-240-485-105

Query Match 3.0%; Score 50.4; DB 15; Length 15767;

Best Local Similarity 49.3%; Pred. No. 0.55;

Matches 132; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 33 TGTGATAAAATTTATAGTTAGTACTTCTTGTAAATCAATTCAGAGGATGAAAAAAAT 92

Db 14988 TATTTTGAGATGGAAGTTAAATTTAAATTTTGAAGTTAAAGTTAAATTTATTTT 15047

QY 93 CGCAAGAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTAATACAAATAT 152

Db 15048 TGTGAAGAATATTTTATAGATAATGTTGTATATTTTATATTTTATATACGAAAGTAT 15107

QY 153 GTAAGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCAATAATGCGCTAAA 212

Db 15108 GTAAATGTTGATTTGTTTATTTTGGTGAATGTTTATTAAGTCAATTTGTTGTTTAAA 15167

QY 213 TTAATAATTCATTTCTATTTTCCCTTAAGAAAAAAACAGCCCATTAAGGACCATTAGAA 272

Db 15168 TTGAATTTAAATATTTTATTTTAAATTTAAATTTGTAATAATTAAGTGTATTAGAG 15227

QY 273 TGGCGCTGCTCCATTTTATATATA 300

Db 15228 AAGTTTATATTTATTTTGTGTTTTTA 15255

RESULT 14

US-10-239-676-206/c
; Sequence 206, Application US/10239676
; Publication No. US20030082609A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation

; FILE REFERENCE: 5013.1003

; CURRENT APPLICATION NUMBER: US/10/239,676

; CURRENT FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: PCT/EP01/03968

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-04-06

; 2000-04-06

; 2000-04-07

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 228

; SEQ ID NO 206

; LENGTH: 17142

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-239-676-206

Query Match 3.0%; Score 50.2; DB 15; Length 17142;

Best Local Similarity 50.6%; Pred. No. 0.65;

```
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 17 ATACCAAGTCATCTACATGTGATAAAATTTATAGTTAGATTACTTCTTGTAATCATTTTCAG 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9510 ATATCAAAATATCTATAAAATAAATCTTTAAACCAAACTTTTATTCAAAAATCTCCA 9451
QY 77 AGGATGAAAAAATAATCGCAAGAACCAATATTTTAAATGAATGATGCAATATACAAAT 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9450 CCTATAATTAAAAACCTATAAAAAAATACTAACTTAAATCTAACCATTCATTCACAAA 9391
QY 137 TTAATTACACAATATGTAAGATTACATGTTTATAGTTTCATAGAAATCAATTTCTAGAGT 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9390 AAAAAGCTTAAAAATTTAAATATCTACCAATAAATTAATTAATATATATCTTTTAAATTT 9331
QY 197 CATATAATAGCCTAAATTAATAATCTTATCTATTTTCTTAAAGAAAAAACAGCCCA 255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9330 CTTAATTATATTAATAAAATTTCCAAAAATAATTACGTAAAAAATAAATATCA 9272

RESULT 15
US-10-311-455-2080/c
; Sequence 2080, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2080
; LENGTH: 17142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2080

Query Match 3.0%; Score 50.2; DB 15; Length 17142;
Best Local Similarity 50.6%; Pred. No. 0.65;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 17 ATACCAAGTCATCTACATGTGATAAAATTTATAGTTAGATTACTTCTTGTAATCATTTTCAG 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9510 ATATCAAAATATCTATAAAATAAATCTTTAAACCAAACTTTTATTCAAAAATCTCCA 9451
QY 77 AGGATGAAAAAATAATCGCAAGAACCAATATTTTAAATGAATGATGCAATATACAAAT 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9450 CCTATAATTAAAAACCTATAAAAAAATACTAACTTAAATCTAACCATTCATTCACAAA 9391
QY 137 TTAATTACACAATATGTAAGATTACATGTTTATAGTTTCATAGAAATCAATTTCTAGAGT 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9390 AAAAAGCTTAAAAATTTAAATATCTACCAATAAATTAATTAATATATATCTTTTAAATTT 9331
QY 197 CATATAATAGCCTAAATTAATAATCTTATCTATTTTCTTAAAGAAAAAACAGCCCA 255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
9330 CTTAATTATATTAATAAAATTTCCAAAAATAATTACGTAAAAAATAAATATCA 9272
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:41:59 : Search time 4952 Seconds
(without alignments)
11780.813 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagcgctcaataac.....ctcgttagtttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/PCURS COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US05 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US080 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Match	Query Length	DB ID	Description
1	1658	100.0	100.0	1658	100	US-60-434-242-1
2	1658	100.0	100.0	1658	110	US-60-531-039-16
3	110.4	6.7	385	26	US-09-620-111B-7324	Sequence 1, Appli
4	99.4	6.0	308	18	US-09-304-517A-40406	Sequence 7324, Ap
5	99.4	6.0	308	19	US-09-371-146A-40406	Sequence 40406, A
6	99.4	6.0	308	42	US-09-985-678A-40406	Sequence 40406, A
7	98.4	5.9	396	35	US-09-865-439A-30469	Sequence 30469, A
8	98.4	5.9	396	76	US-60-207-458-74766	Sequence 74766, A
9	98.4	5.9	470	35	US-09-865-439A-27134	Sequence 27134, A
10	98.4	5.9	470	76	US-60-207-458-71431	Sequence 71431, A
11	98.4	5.9	836	46	US-10-155-881-37309	Sequence 37309, A
12	98.4	5.9	836	51	US-10-425-114-7252	Sequence 7252, Ap
13	98.4	5.9	836	51	US-10-425-114A-7252	Sequence 7252, Ap
14	98.4	5.9	836	88	US-60-313-544-973	Sequence 973, App
15	97.4	5.9	861	27	US-09-654-617-273716	Sequence 273716, A
16	97.4	5.9	861	29	US-09-684-016-273716	Sequence 273716, A
17	97.4	5.9	918	51	US-10-425-114-13779	Sequence 13779, A
18	97.4	5.9	918	51	US-10-425-114A-13779	Sequence 13779, A
19	96.8	5.8	315	18	US-09-304-517A-78600	Sequence 78600, A
20	96.8	5.8	315	19	US-09-371-146A-78600	Sequence 78600, A
21	96.8	5.8	315	24	US-09-565-306-3420	Sequence 3420, Ap
22	96.8	5.8	315	42	US-09-985-678-78600	Sequence 78600, A
23	93.8	5.7	204	18	US-09-304-517A-39577	Sequence 39577, A
24	93.8	5.7	204	19	US-09-371-146A-39577	Sequence 39577, A
25	93.8	5.7	204	42	US-09-985-678-39577	Sequence 39577, A
26	91.4	5.5	537	35	US-09-865-439A-48132	Sequence 48132, A
27	91.4	5.5	537	76	US-60-207-458-92395	Sequence 92395, A
28	90.4	5.5	280	18	US-09-303-031A-3994	Sequence 3994, Ap
29	90.4	5.5	280	18	US-09-304-517A-57027	Sequence 57027, A
30	90.4	5.5	280	19	US-09-371-146A-57027	Sequence 57027, A
31	90.4	5.5	280	36	US-09-894-949-3994	Sequence 3994, Ap
32	90.4	5.5	280	36	US-09-894-949A-3994	Sequence 3994, Ap
33	90.4	5.5	280	42	US-09-985-678-57027	Sequence 57027, A
34	83.4	5.0	437	18	US-09-304-517A-79013	Sequence 79013, A
35	83.4	5.0	437	19	US-09-371-146A-79013	Sequence 79013, A
36	83.4	5.0	437	24	US-09-565-306-1774	Sequence 1774, Ap
37	83.4	5.0	437	42	US-09-985-678-79013	Sequence 79013, A
38	81.4	4.9	460	35	US-09-865-439A-43255	Sequence 43255, A
39	81.4	4.9	460	76	US-60-207-458-87528	Sequence 87528, A
40	78.4	4.7	686	35	US-09-865-439A-31938	Sequence 31938, A
41	78.4	4.7	686	76	US-60-207-458-76235	Sequence 76235, A
42	73.4	4.4	316	35	US-09-865-439A-44537	Sequence 44537, A
43	73.4	4.4	316	76	US-60-207-458-88800	Sequence 88800, A
44	71.4	4.3	291	18	US-09-304-517A-39265	Sequence 39265, A
45	71.4	4.3	291	19	US-09-371-146A-39265	Sequence 39265, A

ALIGNMENTS

RESULT 1
 US-60-434-242-1
 ; Sequence 1, Application US/60434242
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology, LLC
 ; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
 ; TITLE OF INVENTION: Thereof
 ; FILE REFERENCE: 38-15(52826)A
 ; CURRENT APPLICATION NUMBER: US/60/434,242
 ; CURRENT FILING DATE: 2002-12-18
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 1658
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-60-434-242-1
 Query Match 100.0%; Score 1658; DB 100; Length 1658;

Db	1141	CAGGCATATAGGGCCACATAGGCCTTCTGTGGCCAGAAAGGTTTCTTAACATACATCT	1200
Qy	1201	TCAGGCGCGGTGCTGGTGTCTTCAACTCTTCCCTCCGCTCCAGAGGGATAAATACG	1260
Db	1201	TCAGGCGCGGTGCTGGTGTCTTCAACTCTTCCCTCCGCTCCAGAGGGATAAATACG	1260
Qy	1261	GGGTCTACGTGTACCGGCGACGATGGTGGCGCGAGGGGCATCTGTCCCGGGGAATGG	1320
Db	1261	GGGTCTACGTGTACCGGCGACGATGGTGGCGCGAGGGGCATCTGTCCCGGGGAATGG	1320
Qy	1321	CGCAGGCCATCGTTCGCACACGACGTCGTACGTGTGGCCCCCGCGAGGGCTCTCTGCACACGC	1380
Db	1321	CGCAGGCCATCGTTCGCACACGACGTCGTACGTGTGGCCCCCGCGAGGGCTCTCTGCACACGC	1380
Qy	1381	GTGCGGCGCGACGGGCGCGAGGACCTTGCATTTGTCGGCTCGGTGACACGTGCCCT	1440
Db	1381	GTGCGGCGCGACGGGCGCGAGGACCTTGCATTTGTCGGCTCGGTGACACGTGCCCT	1440
Qy	1441	GGTCCGGCTTTGCCGGCTTTGCCCTTGGCTCGCGCGACGTGCCGTCTCCCCACCGGA	1500
Db	1441	GGTCCGGCTTTGCCGGCTTTGCCCTTGGCTCGCGCGACGTGCCGTCTCCCCACCGGA	1500
Qy	1501	CGGCTATAAAGCGCGGCGACCTGATTCCTGCATCTCACAAAGCAAGCAGGACGAAC	1560
Db	1501	CGGCTATAAAGCGCGGCGACCTGATTCCTGCATCTCACAAAGCAAGCAGGACGAAC	1560
Qy	1561	AGCCAACACCTTAACCTAAGAACAGTAGTAGTCCCTGTGTACGTCTAGCAACGATCCGG	1620
Db	1561	AGCCAACACCTTAACCTAAGAACAGTAGTAGTCCCTGTGTACGTCTAGCAACGATCCGG	1620
Qy	1621	CCGAGCTCGCCAGTTTTGCTCGTTAGTTGGGACGGCG	1658
Db	1621	CCGAGCTCGCCAGTTTTGCTCGTTAGTTGGGACGGCG	1658

RESULT 3

```

US-09-620-111B-7324
; Sequence 7324, Application US/09620111B
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1070P
; CURRENT APPLICATION NUMBER: US/09/620.111B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9298
; SEQ ID NO 7324
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: 1..385
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc_feature
; LOCATION: 1..385
; OTHER INFORMATION: Ceres Seq. ID 1339942
; US-09-620-111B-7324

```

Query Match	6.7%;	Score 110.4;	DB 26;	Length 385;
Best Local Similarity	99.1%;	Pred. No. 1.2e-14;		
Matches 111: Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1541	ACAAGCAAAAGCAGCAGCAGCAAGCCAAACCTAAACTTAAGAAACAGTAGTAGTCCCTCTGT	1600
Db	1	ACAAGCAAAAGCAGCAGCAGCAAGCCAAACCTAAACTTAAGAAACAGTAGTAGTCCCTCTGT	60
QY	1601	ACGTCCTAGCAACGATCCGCGCGCAGCTGCCAGTTTTGCTCGTTAGTTTCGG	1652
Db	61	ACGTCCTAGCAACGATCCGCGCGCAGCTGCCAGTTTTGCTCGTTAGTTTCGG	112

RESULT 4

```

US-09-304-517A-40406
; Sequence 40406, Application US/09304517A
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)B
; CURRENT APPLICATION NUMBER: US/09/304,517A
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-304-517A-40406

Query Match 6.0%; Score 99.4; DB 18; Length 308;
Best Local Similarity 99.0%; Pred. No. 3.9e-12;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCAAACACCTAAACTAAAGAACAGTAGTAGTCCTCGTGTACGCTAGCAA 1611
DB 1 AGACGCAACAGCCAAACACCTAACTAAGAACAGTAGTAGTCCTCGTGTACGCTAGCAA 60

QY 1612 CGATCCGCGCGAGCTGCCAGTTTGTCTCGTTAGTTTGGG 1652
DB 61 CGATCCGCGCGAGCTGCCAGTTTGTCTCGTTAGTTTGGG 101

RESULT 5
US-09-371-146A-40406
; Sequence 40406, Application US/09371146A
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: ANNOTATED PLANT GENES
; FILE REFERENCE: 38-21(15097)C
; CURRENT APPLICATION NUMBER: US/09/371,146A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 294310
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-371-146A-40406

Query Match 6.0%; Score 99.4; DB 19; Length 308;
Best Local Similarity 99.0%; Pred. No. 3.9e-12;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCCAACACCTAAACTAAAGAACAGTAGTAGTCCCTGTGTACGCTAGCAA 1611
DB 1 AGCAGCAACAGCCCAACACCTAACTAAGAACAGTAGTAGTCCCTGTGTACGCTAGCAA 60

QY 1612 CGATCCGCGCGAGCTGCCAGTTTGTCTCGTTAGTTTGGG 1652
DB 61 CGATCCGCGCGAGCTGCCAGTTTGTCTCGTTAGTTTGGG 101

RESULT 6
US-09-985-678-40406
; Sequence 40406, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517,255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06

```

```

; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-985-678-40406

Query Match
Best Local Similarity 6.0%; Score 99.4; DB 42; Length 308;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAA 1611
DB 1 AGCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAA 60

QY 1612 CGATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
DB 61 CGATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 101

RESULT 7
US-09-865-439A-30469
; Sequence 30469, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 30469
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-052-Pl-K1-G5
US-09-865-439A-30469

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 35; Length 396;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 60

QY 1613 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
DB 61 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 100

RESULT 8
US-60-207-458-74766
; Sequence 74766, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

```

```

; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 74766
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-052-Pl-K1-G5
US-60-207-458-74766

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 76; Length 396;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 60

QY 1613 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
DB 61 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 100

RESULT 9
US-09-865-439A-27134
; Sequence 27134, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 27134
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(470)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3354-008-Pl-K1-B11
US-09-865-439A-27134

Query Match
Best Local Similarity 5.9%; Score 98.4; DB 35; Length 470;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 1612
DB 1 GCAGCAACAGCCAAACCTAACTAAAGAACAGTAGTAGTCCCTGTACGTCTAGCAAC 60

QY 1613 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
DB 61 GATCCGCCGCCAGCTCGCCAGTTTGTCTGTTAGTTTGGG 100

RESULT 10
US-60-207-458-71431
; Sequence 71431, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.

```

```

; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 71431
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-008-PI-KI-B11
US-60-207-458-71431

Query Match      5.9%; Score 98.4; DB 76; Length 470;
Best Local Similarity 99.0%; Pred. No. 7.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAGCAGCCACACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 1612
Db 1 GCAGCAGCAGCCACACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 60

QY 1613 GATCCGCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 100

RESULT 11
US-10-155-881-37309
; Sequence 37309, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 37309
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
US-10-155-881-37309

Query Match      5.9%; Score 98.4; DB 46; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 60

QY 1613 GATCCGCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 100

RESULT 12
US-10-425-114-7252
; Sequence 7252, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
US-10-425-114-7252

Query Match      5.9%; Score 98.4; DB 51; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 60

QY 1613 GATCCGCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 100

RESULT 13
US-10-425-114A-7252
; Sequence 7252, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114A-7252

Query Match      5.9%; Score 98.4; DB 51; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCACCTAAACTAAAGAACAGTAGTGCCTGTGTAGCTCTAGCAAC 60

QY 1613 GATCCGCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 1652
Db 61 GATCCGCGAGCTCGCCAGTTTTCGTTAGTTGGG 100

RESULT 14
US-60-312-544-973
; Sequence 973, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

```

; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 973
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (108)..(449)
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-60-312-544-973

Query Match 5.9%; Score 98.4; DB 88; Length 836;
Best Local Similarity 99.0%; Pred. No. 9.8e-12;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAAC 1612
Db 1 GCAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAAC 60
QY 1613 GATCCGCCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
Db 61 GATCCGCCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 100

RESULT 15

US-09-654-617-273716
; Sequence 273716, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong Annotated Plant Genes
; TITLE OF INVENTION: 38-21(15097)D
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 273716
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-273716

Query Match 5.9%; Score 97.4; DB 27; Length 861;
Best Local Similarity 99.0%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1554 CAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAACG 1613
Db 24 CAGCAACAGCCCAACACCTAACTAAAGAACAGTAGTCCCTGTGTACGCTAGCAACG 83
QY 1614 ATCCGCCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
Db 84 ATCCGCCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 122

Search completed: September 2, 2004, 12:09:01
Job time : 4955 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 2, 2004, 07:52:53 ; Search time 2370 Seconds
(without alignments)
2423.549 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcagggtcacataac.....ctcgtagttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5837357 seqs, 1732150321 residues

Total number of hits satisfying chosen parameters: 11674714

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:
1: /cgn2_6/ptodata/1/pna/PCT NEW COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06 NEW COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07 NEW COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08 NEW COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09 NEW COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US60 NEW COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	7	US-10-732-721-1
C 2	109	6.6	868	6	US-10-425-115-68623
C 3	101.6	6.1	528	6	US-10-425-115-47987
C 4	100.4	6.1	975	6	US-10-425-115-181022
C 5	84.8	5.1	2211	6	US-10-425-115-82158
C 6	71.4	4.3	1034	6	US-10-425-115-181033
C 7	55	3.3	533	9	US-60-579-062-32930
C 8	50.6	3.1	17280	6	US-10-221-714B-498
C 9	50.4	3.0	805	9	US-60-579-062-23067
C 10	48.8	2.9	207542	5	US-09-949-003C-4513
C 11	48.8	2.9	207542	6	US-10-893-315-148
C 12	48.8	2.9	207542	6	US-10-902-387-493
C 13	48.8	2.9	207557	5	US-09-949-003C-13956
C 14	48.8	2.9	207557	6	US-10-893-315-134
C 15	48.8	2.9	207557	6	US-10-902-387-465
C 16	48.8	2.9	215927	9	US-60-582-609-19505
C 17	48.8	2.9	228854	9	US-60-548-091-5633
C 18	48.2	2.9	791	5	US-09-404-520B-13329
C 19	48	2.9	753	9	US-60-579-062-40662
C 20	47.2	2.8	1882	1	PCT-US03-41761-40265
C 21	47.2	2.8	1882	1	PCT-US03-41761-40265
C 22	47.2	2.8	1882	1	PCT-US03-41761-40265
C 23	47.2	2.8	6106	6	US-10-221-714B-152
C 24	47	2.8	688	9	US-60-579-062-40503

25	46.4	2.8	2448	6	US-10-902-441-777	Sequence 777, App
26	46.4	2.8	2448	7	US-10-724-972A-360	Sequence 360, App
27	46.2	2.8	916	9	US-60-579-062-24705	Sequence 24705, A
C 28	46.2	2.8	7040	6	US-10-221-714B-162	Sequence 162, App
C 29	46.2	2.8	61020	6	US-10-221-714B-513	Sequence 513, App
C 30	46	2.8	594	9	US-60-579-062-41308	Sequence 41308, A
C 31	45.8	2.8	29993	6	US-10-482-823-3	Sequence 3, Appli
C 32	45.8	2.8	29993	6	US-10-482-823-5	Sequence 5, Appli
C 33	45.8	2.8	38342	6	US-10-221-714B-472	Sequence 472, App
C 34	45.6	2.8	600	1	PCT-US03-41761-20202	Sequence 20202, A
C 35	45.6	2.8	600	1	PCT-US03-41761-20202	Sequence 20202, A
C 36	45.6	2.8	600	1	PCT-US03-41766A-20202	Sequence 20202, A
C 37	45.4	2.7	3952	6	US-10-221-714B-8	Sequence 8, Appli
C 38	45.4	2.7	5371	6	US-10-221-714B-526	Sequence 526, App
C 39	45.4	2.7	7133	6	US-10-221-714B-109	Sequence 109, App
C 40	45.2	2.7	8711	6	US-10-221-714B-423	Sequence 423, App
C 41	45.2	2.7	8845	6	US-10-221-714B-265	Sequence 265, App
C 42	45.2	2.7	8883	6	US-10-221-714B-488	Sequence 488, App
C 43	45.2	2.7	12426	6	US-10-221-714B-203	Sequence 203, App
C 44	45.2	2.7	77504	6	US-10-918-711-3720	Sequence 3720, Ap
C 45	45.2	2.7	77504	6	US-10-918-754-16803	Sequence 16803, A

ALIGNMENTS

RESULT 1
US-10-732-721-1
; Sequence 1, Application US/10732721
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/10732,721
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match		100.0%	Score 1658;	DB 7;	Length 1658;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches 1658;		Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	GAATTCACGGCTCACAATACCAAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60		
Db	1	GAATTCACGGCTCACAATACCAAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT	60		
Qy	61	CTTGTAATCATTTTCAGAGGTGAAAAAATCCCAAGAAAGCAAAATTTTAAATGAAT	120		
Db	61	CTTGTAATCATTTTCAGAGGTGAAAAAATCCCAAGAAAGCAAAATTTTAAATGAAT	120		
Qy	121	GATGCAATATACAAATTTAATACACAAATATATGAAGATTACATTTTATAGTTTCATAGA	180		
Db	121	GATGCAATATACAAATTTAATACACAAATATATGAAGATTACATTTTATAGTTTCATAGA	180		
Qy	181	AATCAATTTCTAGAGTCATAAATGCTAAATTTAAATTTCTATTTTCTTCTTAAG	240		
Db	181	AATCAATTTCTAGAGTCATAAATGCTAAATTTAAATTTCTATTTTCTTCTTAAG	240		
Qy	241	AAAAAAGAGCCCAATTAAGGACCAATAGATGCGGCGCTGCTCCATTTTATATATA	300		
Db	241	AAAAAAGAGCCCAATTAAGGACCAATAGATGCGGCGCTGCTCCATTTTATATATA	300		
Qy	301	GAGATATGAGTTGTTGCTCACTAGGCGACCTCCACCTCGCATGCGTGTCTTTATTACATG	360		
Db	301	GAGATATGAGTTGTTGCTCACTAGGCGACCTCCACCTCGCATGCGTGTCTTTATTACATG	360		


```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47987
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(528)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14376C.1
US-10-425-115-47987

Query Match
Best Local Similarity 6.1%; Score 101.6; DB 6; Length 528;
Matches 118; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 924 CTTTGGCTCGATCTCCTTAGTAGAGCCAGTCGGTGATAACGTGTTAAGAACCCCTTGT 983
Db 129 CTTTGGCTCGATCTCCTTAGTAGAGCCAGTCGGTGATAACGTGTTAAGAACCCCTTGT 70
QY 984 ACCGAGTGTAGTCCAA-GCCTGCTTAACTGAATAATAGTAGAACACCAAGGACCAAGAT 1042
Db 69 ACCGAGTGTAGTCCAAAGGGCTCTTANCTGAATAATAGTAGTAATAACAGGACCAAAAGT 10
QY 1043 GTAGAGAGG 1051
Db 9 GTAGAGAGG 1

RESULT 4
US-10-425-115-181022
; Sequence 181022, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 181022
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9666C.1
US-10-425-115-181022

Query Match
Best Local Similarity 6.1%; Score 100.4; DB 6; Length 975;
Matches 101; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1551 CAGCAGCAACAGCCAAACCTAACTAAAGACAGTAGTACCTGTGTAGCTAGCA 1610
Db 21 CAGCAGCAACAGCCAAACCTAACTAAAGACAGTAGTACCTGTGTAGCTAGCA 80
QY 1611 ACGATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 1652
Db 81 ACGATCCGCGCGAGCTCGCCAGTTTGTCTGTTAGTTTGGG 122

RESULT 5
US-10-425-115-82158/c
; Sequence 82158, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 82158
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174942C.1
US-10-425-115-82158

Query Match
Best Local Similarity 5.1%; Score 84.8; DB 6; Length 2211;
Matches 100; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1086 CAAAGGTTACATGATATGGGATCTCTCTCTATTATTATAGAC-AAAGCTAGGTTTCAGG 1144
Db 2211 CCAAGGTTACATGATATGGGATCTCTCTCTATTATTATAGAC-AAAGCTAGGTTTCAGG 2152
QY 1145 CATATGGGCCACATAGGCTTCCTGCGCCCAAGAAAGGTTTCTTAACAC 1192
Db 2151 CATATGGGCCACATAGGCTTCCTGCGCCCTAGAAAGGTTTCTTAACAC 2104

RESULT 6
US-10-425-115-181033
; Sequence 181033, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 181033
; LENGTH: 1034
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9667C.1
US-10-425-115-181033

Query Match
Best Local Similarity 4.3%; Score 71.4; DB 6; Length 1034;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTAGTCCCTGTGTACGTCTAGCAACGATCCGCGCCGAGCTCGCCAGTTTTCG 1639
Db 1 GAACAGTAGTAGTCCCTGTGTACGTCTAGCAACGATCCGCGCCGAGCTCGCCAGTTTTCG 60
QY 1640 TCGTTAGTTTGGG 1652
Db 61 TCGTTAGTTTGGG 73

RESULT 7
US-60-579-062-32930
; Sequence 32930, Application US/60579062
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Kovalic, David K.
; APPLICANT: Larosa, Thomas J.
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K.
; APPLICANT: Wu, Wei
```

```

; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: Methods for Genetic Control of Insect Infestations in Plants and
; FILE REFERENCE: Compositions thereof
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 32930
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Diabrotica virgifera
US-60-579-062-32930

Query Match      3.3%; Score 55; DB 9; Length 533;
Best Local Similarity 47.8%; Pred. No. 0.0011;
Matches 160; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 30 ACATGGTAAATTTTATAGTTAGATTACTCTTGTAATCATTTTCAGAGGATGAAAAA 89
DB 54 AAAGAGATGATGTGAATAATAATAATAATAATAATAATAATAATAATAATAATA 113
QY 90 AATCGCAAGAAAGCAATATTTTAAATGAATGATGCAATATACAAATTTTAAATFACACAAT 149
DB 114 ATTTTAAATATATATATATATTTAAATAATAATAATAATAATAATAATAATAATAAT 173
QY 150 TATGTAAGATTACATTTGTTAGTTTCATAGAAATCAATTTCTAGAGTCATAAATGCCT 209
DB 174 TTTTATTTTATAGAGATGATAAAAAATAAGGAATTTATTTTATAGATATATATTAATAAT 233
QY 210 AAATTTAAATCTTATCTTCTTTTCCCTTAAGAAAAAAGCAGCCATTAAGGACCATTA 269
DB 234 AAAATTAATATATATAGAAATTTTAAATAATAATAATAATAATAATAATAATAATAAT 293
QY 270 GAATGCGCGTGTCTCCATTTTATATATATAGATATAGATTGTTGCTCAGTGGCCAC 329
DB 294 ATATTATGGTATTATATATATATATATATATATATATATATATATATATATATATAT 353
QY 330 TCACCTCGCATGCGTGTCTTTTATACATGAAAA 364
DB 354 TAAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 388

RESULT 8
US-10-221-714B-498/c
; Sequence 498, Application US/10221714B
; GENERAL INFORMATION:
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: tumor suppressor genes and oncogenes
; CURRENT APPLICATION NUMBER: US/10/221.714B
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 498
; LENGTH: 17280
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714B-498

```

```

Query Match      3.1%; Score 50.6; DB 6; Length 17280;
Best Local Similarity 52.1%; Pred. No. 0.073;
Matches 113; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 38 TAAATTTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGGATGAAAAAATCGCAA 97
DB 5424 TAAATTAACAATAATAAATTTCACTCAAAAAAATAATAATAATAATAATAATAATAA 5365
QY 98 GAAAGCAAAATATTTTAAATGAATGATGCAATATACAAATTTTAAATFACACAATATGTAAG 157
DB 5364 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 5305
QY 158 ATTACATGTTTGTAGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCTTAAATAAA 217
DB 5304 AAAAAACATAAATACTCTTCACTTAAATTTTACCATTCTTTAATAATCTTATCTAAAT 5245
QY 218 ATTCTATTCTATTTTCTTAAAGAAAAAAGCAGCC 254
DB 5244 CAACATATACATAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 5208

RESULT 9
US-60-579-062-23067
; Sequence 23067, Application US/60579062
; GENERAL INFORMATION:
; APPLICANT: Baum, James A
; APPLICANT: Kovalic, David K
; APPLICANT: Larosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munyikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: Methods for Genetic Control of Insect Infestations in Plants and
; FILE REFERENCE: Compositions thereof
; CURRENT APPLICATION NUMBER: US/60/579,062
; CURRENT FILING DATE: 2004-06-11
; NUMBER OF SEQ ID NOS: 41445
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 23067
; LENGTH: 805
; TYPE: DNA
; ORGANISM: Diabrotica virgifera
US-60-579-062-23067

Query Match      3.0%; Score 50.4; DB 9; Length 805;
Best Local Similarity 46.6%; Pred. No. 0.02;
Matches 194; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

QY 30 ACATGGTAAATTTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGGATGAAAAA 89
DB 202 AAAAAATCACAATCTTATTGTAACACAAATATTTTACATCATTTTGTGTAAGAAATGC 261
QY 90 AATCGCAAGAAAGCAATATTTTAAATGAATGATGCAATATACAAATTTTAAATFACACAAT 149
DB 262 ATTCGCTTTTAAATATATATATATATATATATATATATATATATATATATATATATAT 321
QY 150 TATGTAAGATTACATTTGTTTGTAGTTTCATAGAAATFCA-ATTTCTAGAGTCATAAATGCC 208
DB 322 TATTTATAATTAATAATATTTTATATATATATATATATATATATATATATATATATATAT 381
QY 209 TAAATTAATAATCTTATCTATTTTCTTAAAGAAAAAAGCAGCCATTAAGGACCAT 268
DB 382 AATAAAAAATTTTATTAACATTTACTTTTAAATAATAATAATAATAATAATAATAATA 441
QY 269 AGAATGCGCGTGTCTCCATTTTATATATATAGAGATATAGATTGTTGCTCAGTGGCCA 328
DB 442 TTAACATATGCTATCATTTATATATATATATATATATATATATATATATATATATATAT 501

```



```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(207542)
; OTHER INFORMATION: n = A,T,C or G
US-10-902-387-493

Query Match      2.9%; Score 48.8; DB 6; Length 207542;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 40 AATTTTATAGTTAGATTACTTCTTGTAATCAATTCAGAGGATCAAAAAAATCGCAAGA 99
Db 168424 AGTATTTTACATGATGATGCTAAATATTAGCTTAAATTAAGAAACTAAGTTAATA 168365

QY 100 AAGCAAAATATTTTAAATGAATGATGCAATATATACAAATTTTAAATTTACAAATTTATGTAAGAT 159
Db 168364 AAAACCTGACAAATTTCTGGAATTTTACACTGGATAAATTTTATGAAATTTGATATGAGATAT 168305

QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCCTTAATTTAAAT 219
Db 168304 TGAATATTTAAATTTGCAAAATTTTAAATTTTGCAGATTAAGAAACAATAATGTTTCTTCACATGACAT 168245

QY 220 TCTATTCTATTTTCTTAAAGAAAAAAGAGCCCATTAAGGACCATTAAGATGCGCGC 279
Db 168244 ACTATGCAATTTTAAATTTTGCAGATTAAGAAACAATAATGTTTCTTCACATGACCA 168185

QY 280 TGCTCCATTTTATATATAT 299
Db 168184 TGTAAATTTTCAATTTCAAT 168165

RESULT 14
US-10-893-315-134/c
; Sequence 134, Application US/10893315
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 134
; LENGTH: 207557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(207557)
; OTHER INFORMATION: n = A,T,C or G
US-10-893-315-134

Query Match      2.9%; Score 48.8; DB 6; Length 207557;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 40 AATTTTATAGTTAGATTACTTCTTGTAATCAATTCAGAGGATCAAAAAAATCGCAAGA 99
Db 168424 AGTATTTTACATGATGATGCTAAATATTAGCTTAAATTAAGAAACTAAGTTAATA 168365

QY 100 AAGCAAAATATTTTAAATGAATGATGCAATATATACAAATTTTAAATTTACAAATTTATGTAAGAT 159
Db 168364 AAAACCTGACAAATTTCTGGAATTTTACACTGGATAAATTTTATGAAATTTGATATGAGATAT 168305

QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATAATAATGCCTTAATTTAAAT 219
Db 168304 TGAATATTTAAATTTGCAAAATTTTAAATTTTGCAGATTAAGAAACAATAATGTTTCTTCACATGACAT 168245

QY 220 TCTATTCTATTTTCTTAAAGAAAAAAGAGCCCATTAAGGACCATTAAGATGCGCGC 279
Db 168244 ACTATGCAATTTTAAATTTTGCAGATTAAGAAACAATAATGTTTCTTCACATGACCA 168185

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2975)..(2994)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (18838)..(18906)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33239)..(35521)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45489)..(45960)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (125952)..(131289)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (158403)..(158422)

```

QY 280 TGCTCCATTTTATATAT 299
||| ||| ||| |||
Db 168184 TGTTAATTATTCATTTCAT 168165

RESULT 15

US-10-902-387-465/c
; Sequence 465, Application US/10902387
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH CARDIOVASCULAR DISEASE AND LIPID DISORDERS, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL000787
; CURRENT APPLICATION NUMBER: US/10/902,387
; CURRENT FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/231,399
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 6404
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 465
; LENGTH: 207557
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(207557)
; OTHER INFORMATION: n = A,T,C or G
US-10-902-387-465

Query Match 2.9%; Score 48.8; DB 6; Length 207557;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 128; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 40 AATTTTATAGTTAGATTACTTCTTGTAATCATTTTCAGAGGATGAAAAAATCGCAAGA 99
Db 168424 AGTATTTTACATGATTAGATGCTAAATATTAGCTTAAAAATAGAAAACTAAGTTAATA 168365
QY 100 AAGCAAATATTTTAAATGAATGATGCAATATACAAATTTAATTACACAAATTATGTAAGAT 159
Db 168364 AAACCTGACAAATTTGTGAAATTTACACTGGATAAATTTATTGAAATTTGATATGAGATAT 168305
QY 160 TACATTTGTTTATGTTTCATAGAAATCAATTTCTAGAGTCATATAATGCTAAATTTAAAT 219
Db 168304 TGAATATTTAATTTGCAAAATTTTAATATATAGATACACATTTTATCTGCATATGACAT 168245
QY 220 TCTATTTCTATTTTCTTAAAGAAAAAAGCCCATTTAAGGACCATTAGAATCGCGC 279
Db 168244 ACTATGCAATTTTAAATTTGCAATAAGAAAAACAATAATGTTTCTTTCATGACCA 168185
QY 280 TGCTCCATTTTATATAT 299
||| ||| ||| |||
Db 168184 TGTTAATTATTCATTTCAT 168165

Search completed: September 2, 2004, 12:48:44
Job time : 2374 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 07:29:38 : Search time 3928 Seconds
(without alignments)
12604.760 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggctcaataac.....ctcgtagttgggacggcg 1658

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
C	1	751	45.3	779	28	CC418961	CC418961 PUEDX81TD
C	2	719.2	43.4	754	29	CG146202	CG146202 PUCV147B
C	3	694.8	41.9	728	29	CG152000	CG152000 PUIFT17TD
C	4	674	40.7	938	29	CG146205	CG146205 PUCV147D

5	482.8	22.1	498	29	CG151999	CG151999 PUIFT17TB	
6	375.4	22.6	862	29	CG175932	CG175932 PUFY251TB	
C	7	373	949	28	BZ705097	BZ705097 PUCB143TD	
C	8	363.4	21.9	949	29	CG041829	CG041829 PUFW117TB
C	9	361.8	21.8	373	29	CG654144	CG654144 QGUDK20TV
C	10	358.8	21.6	879	29	CG118361	CG118361 PUFYU43TB
C	11	343.8	20.7	985	29	CG175933	CG175933 PUFY251TD
C	12	317.4	19.1	776	29	CG113123	CG113123 PUGY13TD
C	13	262.8	15.9	993	29	CC463184	CC463184 ZMMBB034
C	14	250.4	15.1	844	29	CG096594	CG096594 PUFUD53TB
C	15	250.2	15.1	914	29	CG124360	CG124360 PUFUD11TB
C	16	212.4	12.8	728	29	CG164746	CG164746 PUIU51TB
C	17	205.2	12.4	665	28	CC443358	CC443358 PUDIU72TD
C	18	194.4	11.7	775	28	BZ961663	BZ961663 PURN60TD
C	19	194.4	11.7	826	28	BZ988278	BZ988278 PURN60TD
C	20	192.2	11.6	732	28	CC352433	CC352433 PUNHX59TD
C	21	191.4	11.5	786	29	CG062076	CG062076 PUPDP96TD
C	22	191.4	11.5	858	29	CG062075	CG062075 PUPDP96TB
C	23	188	11.3	925	29	CG096596	CG096596 PUFUD53TD
C	24	175.8	10.6	884	28	CC006284	CC006284 PUDIN46TD
C	25	172.2	10.4	446	29	CG193725	CG193725 PUDJ37TD
C	26	170.8	10.3	438	29	CG193723	CG193723 PUDJ37TB
C	27	165	10.0	1032	29	CG046479	CG046479 PUFV38TD
C	28	164.2	9.9	955	29	CG097258	CG097258 PUFWE84TB
C	29	138.4	8.3	656	29	CG050059	CG050059 PUFUA36TB
C	30	138.4	8.3	656	29	CG050060	CG050060 PUFUA36TD
C	31	127.6	7.7	866	29	CG118362	CG118362 PUFVU43TD
C	32	120.8	7.3	695	29	CC752611	CC752611 ZMMBB013
C	33	120.8	7.3	740	29	CG041830	CG041830 PUFW117TD
C	34	112	6.8	949	29	CG266345	CG266345 OG1AE48TV
C	35	110.6	6.7	754	28	BZ882297	BZ882297 PUGAG16TD
C	36	110.6	6.7	807	28	BZ882291	BZ882291 PUGAG16TB
C	37	107.2	6.5	656	28	BH872036	BH872036 hm38d06.b
C	38	107	6.5	912	29	CG833955	CG833955 ZMMBB014
C	39	102.2	6.2	816	28	CC396173	CC396173 PUHLL28TD
C	40	101.6	6.1	528	10	AW216120	AW216120 687029C03
C	41	96.2	5.8	794	28	BZ991151	BZ991151 PUSCF15TD
C	42	95.6	5.8	873	29	CG102130	CG102130 PUKCN52TD
C	43	95.4	5.8	523	28	BZ341248	BZ341248 ic44b07.g
C	44	92.2	5.6	870	11	AY106704	AY106704 Zea mays
C	45	89.4	5.4	871	28	CC366231	CC366231 PUKKX92TB

ALIGNMENTS

RESULT 1	CC418961	PUEDX81TD	ZM_0.6_1.0_KB	Zea mays	genomic clone ZMMBTA230M18,	779 bp	DNA	linear	GSS 19-MAY-2003
LOCUS	CC418961/c								
DEFINITION	CC418961								
ACCESSION	CC418961								
VERSION	CC418961.1								
KEYWORDS	GSS.								
SOURCE	Zea mays								
ORGANISM	Zea mays								
REFERENCE	1 (bases 1 to 779)								
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.								
TITLE	Maize Genomics Consortium								
JOURNAL	Unpublished (2003)								
COMMENT	Contact: Cathy Whitelaw								
	TICK								
	9712 Medical Center Drive, Rockville, MD 20850, USA								
	Tel: 301-838-5843								
	Fax: 301-838-0208								
	Email: whitelaw@tigr.org								
	Seq primer: TF								
	Class: sheared ends.								

FEATURES		Location/Qualifiers		Accession	
source	1..779	genomic survey sequence.		CG146202	
	/organism="Zea mays"		CG146202.1		
	/mol_type="genomic DNA"		GI:34036985		
	/strain="B73"		GSS.		
	/db_xref="taxon:4577"		Zea mays		
REFERENCE	/clone="ZMMBPa230M18"		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	/clone_lib="ZM 0.6 1.0 KB"		Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		
	/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high		clade; Panicoideae; Andropogoneae; Zea.		
	CoT selected genomic DNA library"		1 (bases 1 to 754)		
			Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,		
ORIGIN	45.3%; Score 751; DB 28; Length 779;		Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and		
	Query Match 43.4%; Score 719.2; DB 29; Length 754;		Bennetzen,J.		
	Best Local Similarity 98.6%; Pred. No. 1.4e-157;		Maize Genomics Consortium		
	Matches 768; Conservative 0; Mismatches 10; Indels 1; Gaps 1;		Unpublished (2003)		
			Other_GSSs: PUUCV14TD		
QY	62 TTGTAATCATTTTCAGAGGATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATG 121		Contact: Cathy Whitelaw		
	779 TTGTAATCATTTTCAGAGGATGAAAAAATCGCAAGAAAGCAAAATATTTTAAATGAATG 720		TIGR		
	122 ATGCAATATACAAATTTAATTACACAAATGTAAGATTACATTTAGTTTCATAGAA 181		9712 Medical Center Drive, Rockville, MD 20850, USA		
	719 ATGCAATATACAAATTTAATTACACAAATGTAAGATTACATTTAGTTTCATAGAA 660		Tel: 301-838-5843		
	182 ATCAATTTCTAGAGTCATATAATGCTAAATTAATAATCTATTCTATTCTTCTTAAGA 241		Fax: 301-838-0208		
DB	659 ATCAATTTCTAGAGTCATATAATGCTAAATTAATAATCTATTCTATTCTTCTTAAGA 600		Email: whitelaw@tigr.org		
	242 AAAAAACAGCCCATTAAGGACCATTAGAATCGCGCTGCTCCA-TTTTATTATATATA 300		Seq primer: TR		
	599 AAAAAAACAGCCCATTAAGGACCATTAGAATCGCGCTGCTCCA-TTTTATTATATA 540		Class: Sheared ends.		
	301 GAGATATGAGTTGTTGCTCACTAGGCACATCCACCTCGCATGCTGTTCTTATTACATG 360		Location/Qualifiers		
	539 GAGATATGAGTTGTTGCTCACTAGGCACATCCACCTTCGATGCGTGTCTTATTACATG 480		1..754		
QY	361 AAAAAATGAGCCGCTGCTTTTGCATTAAGGGTCTAGAGATGTACAGTGTCAACGCTTCGT 420		/organism="Zea mays"		
	479 AAAAAATGAGCCGCTGCTTTTGCATTAAGGGCTAGAGATGTACAGTGTCAACGCTTCGT 420		/mol_type="genomic DNA"		
	421 ATTCACTGCACGAGTAATTTTTTTTAAATGACGAGGCGCTTTGTATTCTTCAGCAGCT 480		/strain="B73"		
	419 ATTCACCTGCACGAGTAATTTTTTTTAAATGACGAGGCGCTTTGTATTCTTCAGCAGCT 360		/db_xref="taxon:4577"		
	481 CTCGGCGCCGCTGCTTCGTGTACGGCAGTGCACGAGCAGCAGCAGCTCTACGTGCCACG 540		/clone="ZMMBTA0639C04"		
DB	359 CTCGGCGCCGCTGCTTCGTGTACGGCAGTGCACGAGGCAGCAGCTCTACGTGCCACG 300		/clone_lib="ZM 0.6 1.0 KB"		
	541 TGCCTGCGTGCCTGCCACCGTGATGCAAGGCTTGTCTTGTGCGTGCCTGCGTGTGGC 600		/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high		
	299 TGCCTGCGTGCCTGCCACCGTGATGCAAGGCTTGTCTTGTGCGTGCCTGCGTGTGGC 240		CoT selected genomic DNA library"		
	601 AGCGACGAGTTGAAGCTACGCGCGTGGCTGCTTGCCTGCGCGCTGCTAGCAGCGGCAC 660		Query Match 43.4%; Score 719.2; DB 29; Length 754;		
	239 AGCGACGAGTTGAAGCTACGCGCGTGGCTGCTTGCCTGCGCGGCTGCTAGCAGCGGCAC 180		Best Local Similarity 98.9%; Pred. No. 1.9e-150;		
QY	661 GACGTCAACGCTGATGCTAGCGTTGCTACTTTTGTCTTTTGTCTGCTGCTGCGATG 720		Matches 746; Conservative 0; Mismatches 3; Indels 5; Gaps 2;		
	179 GACGTCAACGCTGATGCTAGCGTTGCTACTTTTGTCTTTTGTCTGCTGCTGCGATG 120		QY		
	721 ATGCAGTAAATAAACGTTGTATCTCATGTGATGATCGACGCTGTGCTGCGCATTTCAA 780		Db		
	119 ATGCAGTAAATAAACGTTGTATCTCATGTGATGATCGACGCTGTGCTGCGCATTTCAA 60		QY		
	781 ATCAGTAAACATCATGTCATCCCCCATCGACTTCGTGTCGCGCATTAATCTGCTAGG 839		Db		
DB	59 ATCAGTAAACATCATGTCATCCCCCATCGACTTCGTGTCGCGCATTAATCTGCTAGG 1		QY		
			Query Match 43.4%; Score 719.2; DB 29; Length 754;		
			Best Local Similarity 98.9%; Pred. No. 1.9e-150;		
			Matches 746; Conservative 0; Mismatches 3; Indels 5; Gaps 2;		
			QY		
QY	867 AATAGAGAGAAAGATAATATATATACATAAATAATGAGTTCAACATCTTTCTTAATACCTT 926		Db		
	754 AATAGAGAGAAAGATAATATATATACATAAATAATGAGTTCAACATCTTTCTTAATACCTT 695		QY		
	927 TTGCTCGATCTCTTTAGTAGAGCCA CGTGC GTGATA CGTGT TTAAGAACCCCTTTGTACC 986		Db		
	694 TTGCTCGATCTCTTTAGTAGAGCCA CGTGC GTGATA CGTGT TTAAGAACCCCTTTGTACC 635		QY		
	987 GAGTGTAGTCCAA -GGCTGCTTAAGTGAATAATAGTAGTAACACAGGACAGAGTGTA 1045		Db		
DB	634 GAGTGTAGTCCAAAGGCTGCTTAACTGAATAATAGTAGTAACACAGGACAGAGTGTA 575		QY		
	1046 GAGAGAACTGATCTTTTGTACTATATGTTGCTGCTCTCCAAAGGTTACATGATATGGG 1105		Db		
	574 GAGAGAACTGATCTTTTGTACTATATGTTGCTGCTCTCCAAAGGTTACATGATATGGG 515		QY		
	1106 GATCTCTCTCTATTTATAGACAAAACTAGGTTTTCAGGCATATGAGGCACATAGGCGCTT 1165		Db		
	514 GATCTCTCTCTATTTATAGACAAAACTAGGTTTTCAGGCATATGAGGCACATAGGCGCTT 455		QY		
QY	1166 CTGGGCCCAAGAAAGTTTCTTAACTACCATCTTCCAGCGCGGTGCTGCGTGTCTCA 1225		Db		
	454 CTGGGCCCAAGAAAGTTTCTTAACTACCATCTTCCAGCGCGGTGCTGCGTGTCTCA 395		QY		
	1226 ACTCTCTCTCTCGCTCCAGAGCGG ----ATAAATACGGGGTCTACGGTGCACGCGCAC 1281		Db		
	394 ACTCTCTCTCTCGCTCCAGAGCGGATAAATAAATACGGGGTCTACGGTGCACGCGCAC 335		QY		
	1282 GCATGTGCGCGAGGGGATCTCTGCGCGGGGAATGCGCGCAGGCGCATCGCTCGACACG 1341		Db		
DB	334 GCATGTGCGCGAGGGGATCTCTGCGCGGGGAATGCGCGCAGGCGCATCGCTCGACACG 275		QY		
	1342 CACGTACGTGTGTCGCCCGCGAGGCGCTCTCTCGACACACGCTTCGCGCGCGCGCGCA 1401		Db		
			QY		
			Query Match 43.4%; Score 719.2; DB 29; Length 754;		
			Best Local Similarity 98.9%; Pred. No. 1.9e-150;		
			Matches 746; Conservative 0; Mismatches 3; Indels 5; Gaps 2;		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
			QY		
			Db		
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			
		QY			
		Db			


```
Db 274 CACGTACGTGTGCGCCCGCGAGGCTCTCGACACGGCTCGCGCGCACGGCGCGCA 215
QY 1402 GGGACCTTGCAATTCGTCCGGTCGGTGACACAGTCCCTGGTCCGGCTTGCCTTGC 1461
Db 214 GGGACCTTGCAATTCGTCCGGTCGGTGACACAGTCCCTGGTCCGGCTTGCCTTGC 155
QY 1462 CTTGCGCTCGCGCGAGCGTGGCGTCTCCCGACCGAGCGCTATAAAAAGCGCGCGC 1521
Db 154 CTTGCGCTCGCGCGAGCGTGGCGTCTCCCGACCGAGCGCTATAAAAAGCGCGCGC 95
QY 1522 ACCTGATCCTCATCTCACAAAGCAAGCAGCAGCAACAGCAACCTAACTAAAGA 1581
Db 94 ACCTGATCCTCATCTCACAAAGCAAGCAGCAGCAACAGCAACCTAACTAAAGA 35
QY 1582 ACAGTAGTAGTCCCTGTGTAGTCTAGCAACGAT 1615
Db 34 ACAGTAGTAGTCCCTGTGTAGTCTAGCAACGAT 1

RESULT 3
CG152000/c
LOCUS PU1F117D ZM.0.6.1.0 KB Zea mays genomic clone ZMMBTA0573C10,
DEFINITION genomic survey sequence.
ACCESSION CG152000
VERSION CG152000.1 GI:34042783
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 728)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PU1F117B
Contact: Cathy Whitelaw
TIOR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: 1F
Class: sheared ends.
Location/Qualifiers
1..728
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA0573C10"
/clone_lib="ZM.0.6.1.0 KB"
/note="Vector: pCR4-TORO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 41.9%; Score 694.8; DB 29; Length 728;
Best Local Similarity 99.0%; Pred. No. 5.5e-145;
Matches 721; Conservative 0; Mismatches 2; Indels 5; Gaps 2;

QY 601 AGCGACGAGTTGAAGCTACGCGCGTGGCTGCTTGCCTGCGCGCTAGCAGCGGAC 660
Db 728 AGCGACGAGTTGAAGCTACGCGCGTGGCTGCTTGCCTGCGCGCTAGCAGCGGAC 669
QY 661 GACGTCGACGCTGATGCTAGCGTCTACTTTGTTTGTCTCTGCTGCGGATG 720
Db 668 GACGTCGACGCTGATGCTAGCGTCTACTTTGTTTGTCTCTGCTGCGGATG 609
QY 721 ATGCAAGTAAATAAACGTTGTATCTCATGTGATTGATCGAGCTGTGTCGGGATTTC 780
```

```
Db 608 ATGCAAGTAAATAAACGTTGTATCTCATGTGATTGATCGAGCTGTGTCGGGATTTC 549
QY 781 ATCAGTAAACAATCACAATGTCATCCCATCGACTTCGTGTCGCGGATAAATCTGTAGG 840
Db 548 ATCAGTAAACAATCACAATGTCATCCCATCGACTTCGTGTCGCGGATAAATCTGTAGG 489
QY 841 ATTGTGCTTTTCATGTATGTCTTGAATAGAGAGAAAGAATAATATTACATAAATAT 900
Db 488 ATTGTGCTTTTCATGTATGTCTTGAATAGAGAGAAAGAATAATATTACATAAATAT 429
QY 901 GAGTTCAACATCTTCTTAAATACCTTTTCTCGATCTCCTTAGTAGAGCACTGCGTGA 960
Db 428 GAGTTCAACATCTTCTTAAATACCTTTTCTCGATCTCCTTAGTAGAGCACTGCGTGA 369
QY 961 TAACTGTTTAAGAACCCCTTGTACCGAGTGTAGTCCAA-GGCTGCTTAACTGAATAATA 1019
Db 368 TAACTGTTTAAGAACCCCTTGTACCGAGTGTAGTCCAAAGGGCTGCTTAACTGAATAATA 309
QY 1020 GTAGAACACAGGACACAGAGTGTAGAGAGAACTGATCTTTGTTACTATATGTTGCT 1079
Db 308 GTAGAACACAGGACACAGAGTGTAGAGAGAACTGATCTTTGTTACTATATGTTGCT 249
QY 1080 GCTCTCCAAAGCTTACATGATATGGGATCTCCTCTCTATTATAGACAAACTAGGTT 1139
Db 248 GCTCTCCAAAGCTTACATGATATGGGATCTCCTCTCTATTATAGACAAACTAGGTT 189
QY 1140 TCAGGCATATGGGCCACATAGGCTTCTTGGCCCAAGAAAGGTTTCTTAACTACCATC 1199
Db 188 TCAGGCATATGGGCCACATAGGCTTCTTGGCCCAAGAAAGGTTTCTTAACTACCATC 129
QY 1200 TTCAGAGGCGGTGCTGGTGGCTTCACTCTTCTTCCCTGCGCTCCAGAGCGG----ATAA 1255
Db 128 TTCAGAGGCGGTGCTGGTGGCTTCACTCTTCTTCCCTGCGCTCCAGAGCGGATAATAA 69
QY 1256 ATACGGGCTGTACGCTGCACGCGCACGCGATGTGCGCGAGGGGCTCTCTGCCCGGGA 1315
Db 68 ATACGGGCTGTACGCTGCACGCGCACGCGATGTGCGCGAGGGGCTCTCTGCCCGGGA 9
QY 1316 ATGCGCGC 1323
Db 8 ATGCGCGC 1

RESULT 4
CG146205
LOCUS PUUCV14TD ZM.0.6.1.0 KB Zea mays genomic clone ZMMBTA0639C04,
DEFINITION genomic survey sequence.
ACCESSION CG146205
VERSION CG146205.1 GI:34036988
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 938)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUJCV14TB
Contact: Cathy Whitelaw
TIOR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: 1F
Class: sheared ends.
Location/Qualifiers
```

```
source
1. .938
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0639C04"
/clone_lib="ZM 0.6 1.0 KB"
/notes="vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 40.7%; Score 674; DB 29; Length 938;
Best Local Similarity 98.3%; Pred. No. 2.5e-140; Indels 2; Gaps 2;
Matches 702; Conservative 0; Mismatches 10;

QY 1 GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTTATAGTTAGTACTT 60
DB 225 GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTTATAGTTAGTACTT 284
QY 61 CTTGTAATCATTTTCAGAGGATGAAGAAAAAATCGCAGAAAGCAAAATATTTAAATGAAT 120
DB 285 CTTGTAATCATTTTCAGAGGATGAAGAAAAAATCGCAGAAAGCAAAATATTTAAATGAAT 344
QY 121 GATGCAATATACAAATTTAATTAACAAATATATGTAAAGATTACATTTGTTAGTTTCATAGA 180
DB 345 GATGCAATATACAAATTTAATTAACAAATATATGTAAAGATTACATTTGTTAGTTTCATAGA 404
QY 181 AATCAATTTCTAGAGTCATTAATATGCTAAATTAATAATTTCTATTTCTTTTCTTAAAG 240
DB 405 AATCAATTTCTAGAGTCATTAATATGCTAAATTAATAATTTCTATTTCTTTTCTTAAAG 464
QY 241 AAAAAAAGACCCATTAAAGGACCATTAAGATCGCGCTGCTCCA-TTTTTTATATAT 299
DB 465 AAAAAAAGACCCATTAAAGGACCATTAAGATCGCGCTGCTCCA-TTTTTTATATAT 524
QY 300 AGAGATATGATGTTGCTCACTAGGCCACTCCACCTTCGCATGCGTCTTTTATTACAT 359
DB 525 AGAGATATGATGTTGCTCACTAGGCCACTCCACCTTCGCATGCGTCTTTTATTACAT 584
QY 360 GAAAAAATAGCGCTGCTTTTGGATTAAGGCTTAGAGATGTACAGTGTCAACGCTTCG 419
DB 585 GAAAAAATAGCGCTGCTTTTGGATTAAGGCTTAGAGATGTACAGTGTCAACGCTTCG 644
QY 420 TATTCTAGCAGCAGTAAATTTTTTTAATGCAAGGAGCGCTTTGATTCTCAGCAGC 479
DB 645 TATTCTAGCAGCAGTAAATTTTTTTAATGCAAGGAGCGCTTTGATTCTCAGCAGC 704
QY 480 TCTGCGCGCGCTGCTTTCGTGACGCGAGTGACGAGGCAGCAGCTCTACGTGCCAC 539
DB 705 TCTGCGCGCGCTGCTTTCGTGACGCGAGTGACGAGGCAGCAGCTCTACGTGCCAC 764
QY 540 GTGCTGGCTGCTGCCACCGTGTATGGCAAGGCTTGTCTGCTGCGTGGCTGGTGG 599
DB 765 GTGCTGGCTGCTGCCACCGTGTATGGCAAGGCTTGTCTGCTGCGTGGCTGGTGG 824
QY 600 CAGCGACGAGGTTCAAGCTACGCGGTGGCTGCTGCTGCGCGCTGTAGCACGCGGA 659
DB 825 CAGCGACGAGGTTCAAGCTACGCGGTGGCTGCTGCTGCGCGCTGTAGCACGCGGA 884
QY 660 CGAGCTCGAGCTGATGCTAGCTGTCTAC-TTTTGTGTTTGTGTTTCTCTGTGC 712
DB 885 CGAGCTCGAGCTGATGCTAGCTGTCTAC-TTTTGTGTTTGTGTTTCTCTGTGC 938

RESULT 5
CG151999 498 bp DNA linear GSS 21-AUG-2003
LOCUS PUFT17TB ZM.0.6.1.0 KB Zea mays genomic clone ZMMBTA0573C10,
DEFINITION genomic survey sequence.
ACCESSION CG151999
VERSION CG151999.1 GI:34042782
KEYWORDS GSS.
SOURCE Zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 498)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFT17TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
Location/Qualifiers
1..498
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0573C10"
/clone_lib="ZM 0.6 1.0 KB"
/notes="vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 29.1%; Score 482.8; DB 29; Length 498;
Best Local Similarity 99.4%; Pred. No. 1.8e-97; Indels 1; Gaps 1;
Matches 495; Conservative 0; Mismatches 2;

QY 266 ATTAGAATCGCGCTGCTCCA-TTTTTTATATATATAGATATAGTTTGTCTACTAG 324
DB 1 ATTAGAATCGCGCTGCTCCA-TTTTTTATATATATAGATATAGTTTGTCTACTAG 60
QY 325 GCCACTCCACCTCGCATGGTGTCTTATTATGAAAAAATGAGCGCTGCTTTTGCA 384
DB 61 GCCACTCCACCTCGCATGGTGTCTTATTATGAAAAAATGAGCGCTGCTTTTGCA 120
QY 385 TTAAGGCTCTAGAGATGTACAGTGCACACGCTTCGTATTCACTGCACGAGTAATTTTTT 444
DB 121 TTAAGGCTCTAGAGATGTACAGTGCACACGCTTCGTATTCACTGCACGAGTAATTTTTT 180
QY 445 TTAATGACGAGGCGCTTTGTATTCTCTAGCAGCTCTGCGCGCGCTGCTTGTGTACG 504
DB 181 TTAATGACGAGGCGCTTTGTATTCTCTAGCAGCTCTGCGCGCGCTGCTTGTGTACG 240
QY 505 GCAGTGACGAGGCGAGCAGCTCTAGTGCACACGCTGCTGCTGCCACCGTGA 564
DB 241 GCAGTGACGAGGCGAGCAGCTCTAGTGCACACGCTGCTGCTGCCACCGTGA 300
QY 565 TGGCAAGGCTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
DB 301 TGGCAAGGCTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 625 GTGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
DB 361 GTGCTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 685 GCTACTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 744
DB 421 GCTACTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 480
QY 745 CTCATGTGATTGATCGAC 762
DB 481 CTCATGTGATTGATCGAC 498

RESULT 6
```

```

CG175932
LOCUS          CG175932      862 bp      DNA      linear      GSS 21-AUG-2003
DEFINITION    PUFY251TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMB7a0765106,
               genomic survey sequence.
ACCESSION     CG175932
VERSION       CG175932.1  GI:34066730
KEYWORDS      GSS.
SOURCE        Zea mays
ORGANISM      Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 862)
AUTHORS       Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
               Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
               Bennetzen,J.
TITLE         Maize Genomics Consortium
JOURNAL       Unpublished (2003)
COMMENT       Other GSSs: PUFY251TD
               Contact: Cathy Whitelaw
               TIGR
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301-838-5843
               Fax: 301-838-0208
               Email: whitelaw@tigr.org
               Seq primer: TR
               Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..862
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZMMB7a0765106"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                     Cot selected genomic DNA library"
ORIGIN
Query Match      22.6%; Score 375.4; DB 29; Length 862;
Best Local Similarity 82.4%; Pred. No. 2.1e-73;
Matches 520; Conservative 0; Mismatches 91; Indels 20; Gaps 7;

QY 576 GCTGCTTGGCGTGGTGGCGAGCGAGCTGGAAGTACGCGGCTGGCTGCTG 635
      |||
Db 85 GCTCGTTGACGCTGGTGGTGGTAGCGAGCGAGCGCGGC-ACGACGATGGTACTG 143
      |||
QY 636 CTGCGCGCGTCTAGCAGCGAGCGTGCAGCGCTGATGCTAGCGTGTGCTTCTTTGT 695
      |||
Db 144 CTGTGTGCGTGCAAGCGTGGCACTGAAGCCAGCCCT-----ATAGTTTTT 190
      |||
QY 696 TTTGTTTTCTGTGCTGTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 755
      |||
Db 191 CTTTCTTTTCTATGTTCTTGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 250
      |||
QY 756 GATCGAGCTGTGTCGGCGATTTCAATCAGTAACATCACATGTCATGCCCATCGACTT 815
      |||
Db 251 TCGAGCTGTGTGCGCGATTTCAATCAGTAACATCACATGTCATGCCCATCGACTT 310
      |||
QY 816 CGTGTGCGCGATAAATCTGCT-AGGGATTTGTGCTTTTCATGTATGTGCTTGAATAGAGA 874
      |||
Db 311 CGTGTGAGCAATAATCTACTTCGGGATTTGTGCTTTTCATGTATGTGCTTGAATAGAGA 370
      |||
QY 875 GGAAGAATAATTTA--CATAAATATGAGTTCAACATCTTTCTTAATACCTTTTGCTC 932
      |||
Db 371 GAAAGAACAATATTATACGATAAAATATGAGTTCAACGCTTTCTTAAATACCTTTTGCTC 430
      |||
QY 933 GATCTCCTTAGTAGAGCCAGCGCTGATAACGCTTAAAGACCCCTGTTACCGAGTGT 992
      |||
Db 431 AATCTCCTTAGTAGAGCCAGCGCTGATAAAGTGTAAAGACCCCTGTTACCGAGTAT 490
      |||
QY 993 AGTCCAAGG-CTGCTTAACCTGAATAATAGTAGAACAAGGAGACAAGAGGTGTAGAGG 1051
      |||

```

```

Db 491 AGTCCAAGGACTCTTTAACTGAATATAGTAGAACAACAGGAGACAGAGTGTAGAGG 550
QY 1052 AACTGATTTCTTTGTTACTATATGTTGCTCTCTCAAAGGTTACATGATATGGGATCTC 1111
      |||
Db 551 GACTGATTTCTTTTACTGTATGTTGCTGCTGCTCC-CAGGTTACATGATATGGGATCTC 609
      |||
QY 1112 CTCTCTATTATAGAC-AAAACTAGGTTTCAGGCATATGGCCACATAGGCTTCTCTGG 1170
      |||
Db 610 CTCTCTATTATAGACAAAACCTAAGGTTTCGATATATGGGCCACATGGGCTTCTCTGG 669
      |||
QY 1171 CCCAAGAAAGGTTTCTTAACACTACCATCTT 1201
      |||
Db 670 CCCAATAAAGTTTCTTAACAGCCCAACCT 700
      |||

RESULT 7
BZ705097/c
LOCUS          BZ705097      949 bp      DNA      linear      GSS 19-FEB-2003
DEFINITION    PUCEI43TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMB7a127G13,
               genomic survey sequence.
ACCESSION     BZ705097
VERSION       BZ705097.1  GI:28425505
KEYWORDS      GSS.
SOURCE        Zea mays
ORGANISM      Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 949)
AUTHORS       Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
               Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
               Bennetzen,J.
TITLE         Maize Genomics Consortium
JOURNAL       Unpublished (2003)
COMMENT       Contact: Cathy Whitelaw
               TIGR
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301-838-5843
               Fax: 301-838-0208
               Email: whitelaw@tigr.org
               Seq primer: TP
               Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..949
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone_lib="ZMMB7a127G13"
                     /clone_lib="ZM_0.6_1.0_KB"
                     /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                     Cot selected genomic DNA library"
ORIGIN
Query Match      22.5%; Score 373; DB 28; Length 949;
Best Local Similarity 97.3%; Pred. No. 7.3e-73;
Matches 390; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 GAAATTCAGGCTCAATATCCAGTCACTACATGTGATAAAATTTTATAGTTAGATTACTT 60
      |||
Db 409 GAAATTCAGGCTCAATATCCAGTCACTACATGTGATAAAATTTTATAGTTAGATTACTT 350
      |||
QY 61 CTTGTATCATTTTCAGAGATGAAAAAATTCGCAAGAAAGCAAAATATTTTAAATGAAT 120
      |||
Db 349 CTTGTATCATTTTCAGAGATGAAAAAATTCGCAAGAAAGCAAAATATTTTAAATGAAT 290
      |||
QY 121 GATCAATATCAAAATTTAATTCACATTAATGTAAGATTACATTTAGTTTTCATAGA 180
      |||
Db 289 GATCAATATCAAAATTTAATTCACATTAATGTAAGATTACATTTAGTTTTCATAGA 230
      |||
QY 181 AATCAATTTCTAGAGTCATAATAATGCTTAAATTTAAATTTCTATTCTTTTCTTAAG 240
      |||
Db 229 AATCAATTTCTAGAGTCATAATAATGCTTAAATTTAAATTTCTATTCTTTTCTTAAG 170
      |||

```

```

QY 241 AAAAAAAAAAGCCCATTAAGGACCAATTAGATGCGCGTGTCCCA-TTTTATTATAT 299
Db 169 AAAAAAAAAAGCCCATTAAGGACCAATTAGATGCGCGTGTCCCA-TTTTATTATAT 110
QY 300 AGAGATATGAGTGTGTTGCTCACTAGGCCCACTCCACCTCGCATCGGTGTTCTTTATTACAT 359
Db 109 AGAGATATGAGTGTGTTGCTCACTAGGCCCACTCCACCTTCGATGCGTGTCTTTATTACAT 50
QY 360 GAAAAAATGAGCGCGTGTCTTTTGCATTAAGGCTTAGAGAT 400
Db 49 GAAAAAATGAGCGCGTGTCTTTTGCATTAAGGCTTAGAGAT 9

RESULT 8
CG041829/c
LOCUS
DEFINITION
  CG041829 949 bp DNA linear GSS 19-AUG-2003
  PUFW117B ZM 0.61.0 KB Zea mays genomic clone ZMBMa0748C09,
  genomic survey sequence.
ACCESSION
  CG041829
VERSION
  CG041829.1 GI:33913985
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 949)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
  Bennetzen, J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUFW117D
  Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TR
    Class: sheared ends.
  FEATURES
    source
      1..949
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMBMa0748C09"
        /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
        Cot selected genomic DNA library"
  ORIGIN
    Query Match 21.9%; Score 363.4; DB 29; Length 949;
    Best Local Similarity 84.5%; Pred. No. 1e-70;
    Matches 432; Conservative 0; Mismatches 76; Indels 3; Gaps 2;

QY 700 GTTTCCTCTGTGCTGCGATGTCAGTAAATAAAGCTGTGTATCTCATGTGATC 759
Db 632 GTTGTGTGTGCTCTTTTGGTTATATGATAAAATAAACGTGTGCATATATGTGATTC 573
QY 760 GAGGTGTGCGCGATTTTCAAAATCAGTAACAATACATGTGCATCCCATCGACTTCGTG 819
Db 572 GAGTGTGTGAGGACTTCAAGTCAGCAGCAATACATGTACATCTCCATCGACTTCGTG 513
QY 820 TCGCGGATAAATCTGCTAGGAAATTTGCTTTTCATGTATGTCTTGAATAGAGAGAAA 879
Db 512 CCGCGGATAAATCAGCTTGGGATGTGTATTTCTCGTCATGTCTGTAATAGAAAATAAG 453
QY 880 GAATATATTACAT--AAATATCAGTTCAACATCTTTCTTAATACCTTTTGTCTCATCT 937
Db 452 AACACACTATACATAAGATATATAGTTTCAGCACCTTCTTAGTACCTTTTGTCTCATCT 393

```

```

QY 938 CCTTACTAGAGCACGTGCGTGATACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCC 997
Db 392 TCTTAGTAGAGCACGTGCGTGATACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCC 333
QY 998 AA-GGCTGCTTAACCTGAATAATAGTAGAACAACAAGAGACAAGAGTGTAGAGAGAACTG 1056
Db 332 AAGGGCTGCTTAACCTGAATAATAGTAGAACAACAAGAGACAAGAGTGTAGAGAGAACTG 273
QY 1057 ATCTTTGTATATATGATGCTGCTCTCCAAAGGTTACATGATATGGGATCTCTCTC 1116
Db 272 ATCTTTGTATATATGATGCTGCTCTCCAAAGGTTACATGATATGGGATCTCTCTC 213
QY 1117 TATTTATAGACAAACTAGGCTTTTTCAGGCATATGGGCCACATAGGCTTCTCTGGCCCAAG 1176
Db 212 TATTTATAGACAAACTAGGCTTTTTCAGGCATATGGGCCACATAGGCTTCTCTGGCCCAAG 153
QY 1177 AAAGGTTTCTTAACACTACCATCTTCCAGGC 1207
Db 152 AAAGGTTTCTTAACAGAAACAACATGTAGCC 122

RESULT 9
CC654144
LOCUS
DEFINITION
  CC654144 373 bp DNA linear GSS 19-JUN-2003
  OGUDK20TV ZM 0.71.5 KB Zea mays genomic clone ZMBMa0415C15,
  genomic survey sequence.
ACCESSION
  CC654144
VERSION
  CC654144.1 GI:32057168
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 373)
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
  Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Contact: Cathy Whitelaw
  TIGR
    9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 301-838-5843
    Fax: 301-838-0208
    Email: whitelaw@tigr.org
    Seq primer: TF
    Class: sheared ends.
  FEATURES
    source
      1..373
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMBMa0415C15"
        /clone_lib="ZM 0.7 1.5 KB"
        /note="Vector: pBCSk-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
  ORIGIN
    Query Match 21.8%; Score 361.8; DB 29; Length 373;
    Best Local Similarity 99.5%; Pred. No. 2.5e-70;
    Matches 363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1288 TCGCGCAGGGGATCTCTGCGGGGAATGCGCAGGCGCATCGTCGACACGACGTA 1347
Db 1 TCGCGCAGGGGATCTCTGCGGGGAATGCGCAGGCGCATCGTCGACACGACGTA 60
QY 1348 CGGTGTGCGCCCGCGAGGCTCTCTCGCACACGCGTTCGCGCGCACCGGGCCCGAGGACC 1407
Db 61 CGGTGTGCGCCCGCGAGGCTCTCTCGCACACGCGTTCGCGCGCACCGGGCCCGAGGACC 120

```

```

QY 1408 TTGCAATTGTCGCGGTCCGTCACACAGTCCTGTCGCGCTTGCCTGCTTGCCTGCTG 1467
Db 121 TTGCAATTGTCGCGGTCCGTCACACAGTCCTGTCGCGCTTGCCTGCTTGCCTGCTG 180
QY 1468 CTCGCGCGGACGTCCTCCGCTTCCCCACCGGACGGCTATATAAAGGCGGCGCCACCTGA 1527
Db 181 CTCGCGCGGACGTCCTCCGCTTCCCCACCGGACGGCTATATAAAGGCGGCGCCACCTGA 240
QY 1528 TCCTCCATCTCACACAAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 1587
Db 241 TCCTCCATCTCACACAAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 300
QY 1588 GTAGTCCCTGTGTCAGTCTAGCAAGCATCCGCGCGGAGCTCCGCCAGTTTTCGCTGTTAGT 1647
Db 301 GTAGTCCCTGTGTCAGTCTAGCAAGCATCCGCGCGGAGCTCCGCCAGTTTTCGCTGTTAGT 360
QY 1648 TTGGG 1652
Db 361 TTGGG 365

RESULT 10
CG118361/c
LOCUS
DEFINITION
  PUFVU43TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTa0744H13,
  genomic survey sequence.
ACCESSION
  CG118361
VERSION
  CG118361.1 GI:34001798
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 879)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUFVU43TD
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..879
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBTa0744H13"
    /clone_lib="ZM_0.6-1.0 KB"
    /note="vector: PCR4-ToPO; Site 1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"

ORIGIN
  Query Match 21.6%; Score 358.8; DB 29; Length 879;
  Best Local Similarity 88.2%; Pred. No. 1.1e-69;
  Matches 413; Conservative 0; Mismatches 52; Indels 3; Gaps 2;

QY 727 TAAATAAACGTGTGTATCTCATGTGATTCGACGTGTGTCGCGGATTTCAATCAGT 786
Db 876 TAAATAAACGTGTGTATCTCATGTGATTCGACGTGTGTCGCGGATTTCAATCAGT 817
QY 787 AACATCATGTGCATCCCATCGACTTCGTCGCGGATATAAATCTCTAGGGATTTGT 846
Db 816 AGCAATCATGTACATTCCTCCATCGACTTCGTCGCGGATATAAATCAGCTTGGGATGT 757

```

```

QY 847 GCTTTTCATGTATGCTTGAATAGAGAGAAATATATATACAT--AAATATAGT 904
Db 756 ATTTCCTGTCATGCTGCTGAATAGAAAATAAAGAACACACTATACATAGATATATAGT 697
QY 905 TCAACATCTTTCTTAATACCTTTTCTCGATCTCCTTAGTAGAGCACGTCGCTGATAC 964
Db 696 TCAGACCTTTCTTAGTACCTTTTCTGCTGATCTTTCTTAGTAGAGCACGTCGCTGATAC 637
QY 965 GTGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAA--GCTGCTTAAGTAAATAGTAG 1023
Db 636 GTGTTAAGAACCCCTTTGTTACCGAGTGTAGTCCAAAGGCTGCTTAACGAATAAGTAG 577
QY 1024 AACCAAGGACACAGAGTGTAGAGAGCACTGATCTTTGTTACTATATGCTGCTGTC 1083
Db 576 AACCAAGGACACAGAGTGTAGAGAGCACTGATCTTTGTTACTATATGCTGCTGTC 517
QY 1084 TCCAAAGGTTACATGATATGGGATCTCTCTCTATTTATAGACAAAACCTAGGGTTTCAG 1143
Db 516 TCCAAAGGTTACATGATATGAGATCTCTCTCTATTTATAGACAAAACCTAGGGTTTCAG 457
QY 1144 GCATATGGGCCACATAGGCCCTTCCTGGCCCAAGAAAGGTTTCTTTAACA 1191
Db 456 TCATATGGCGCACATGGGCTTCCTGGCTAAGAAAGGTTTCTTTAACA 409

RESULT 11
CG175933/c
LOCUS
DEFINITION
  PUFVZ51TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTa0765I06,
  genomic survey sequence.
ACCESSION
  CG175933
VERSION
  CG175933.1 GI:34066731
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 985)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUFVZ51TB
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
  Location/Qualifiers
    1..985
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBTa0765I06"
    /clone_lib="ZM_0.6-1.0 KB"
    /note="vector: PCR4-ToPO; Site 1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"

ORIGIN
  Query Match 20.7%; Score 343.8; DB 29; Length 985;
  Best Local Similarity 88.1%; Pred. No. 2.6e-66;
  Matches 431; Conservative 0; Mismatches 52; Indels 6; Gaps 5;

QY 718 ATGATGAGTAAATAAACGTCGTGTATCTCATGTGATTCGACGTGTGTCGCGGATTT 777
Db 982 ATGATGAGTAAATAAACGTCGTGTATCTCATGTGATTCGACGTGTGTCGCGGATTT 923

```



```

/lab_host="E. coli DH10B"
/clone_lib="ZM5Bc"
/note="Vector: pTARAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match      15.9%; Score 262.8; DB 29; Length 993;
Best Local Similarity 73.4%; Pred. No. 3.7e-48;
Matches 496; Conservative 0; Mismatches 142; Indels 38; Gaps 11;

QY 532 GTGCCACGCGCTGGCTGCTGCCACCGCTGATGCGAAGCTTGCTGCTGCGGCGCT 591
Db 731 GCCTGCTGCCACACTCGTCAGCAGTGAAGCGTGGCCCGCCACTCGCTCGTGACCT 672
QY 592 GCGTGGGACGACGAGGTGAAGCTACGGCGGTGCTTGCCTGGCGGCTGCTAG 651
Db 671 TCGTAGGACGCGGTGAGCGCAGTCTACGACGGCTGCTGCTGCTCCATGCTAG 612
QY 652 CAGCGGACGACGTCGACGCTGATGCTAGGCTGCTGCTGCTGCTGCTGCTGCTG 711
Db 611 TACAGCGACT-----ACGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
QY 712 CTGCGCATGATGAGTAATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
Db 564 -----AATGATGAGGAATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
QY 770 GCGATTTCAATCAGTAATCAATGTCATGTCATCCCGCTGCTGCTGCTGCTGCTGCT 829
Db 509 GCGATTTAAAGTCAGCGCAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
QY 830 ATCTGCTAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
Db 450 AATCTGCTAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 885 ATATTACATAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
Db 390 AGATCATAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 332
QY 945 AGAGCCAGCTGGG-----TGATAACGCTGTTAAGAACCCCTGTTTACCAGGTAGTCC 997
Db 331 AGAGCCAGCTGGGTGATACTGATACTGATACTGATACTGATACTGATACTGATACTG 272
QY 998 AA-GGCTGTTAACTGATAAATAGTAGAACACAAGGAGACAAGAGTGTAGAGGAACTG 1056
Db 271 AAGGGCTCCTTAACTGAATAATAGTAGAACACAAGGAGACAAGAGTGTAGAGGAACTG 212
QY 1057 ATCTTTGTTACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
Db 211 ATCTTTTATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154
QY 1117 TATTTATAGACAA-AACTAGGCTTTTACGCAATATGGCCATATGGCCATATGGCCCAA 1175
Db 153 TATTTATAGACAAATAGGCTTTTAAAGTATATGGCCATATGGCCATATGGCCCAA 94
QY 1176 GAAAGTTTCTTAAACA 1191
Db 93 GACAAGATTTCTTAAA 78

RESULT 14
LOCUS CG096594
DEFINITION PUFUD53TB ZM_0.6.1.0 KB Zea mays genomic clone ZM5Bc073J10,
genomic survey sequence.
ACCESSION CG096594
VERSION CG096594.1 GI:33978888
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 844)
REFERENCE
AUTHORS Whitehead,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

```

```

Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUFUD53TD
Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..844
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.6.1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

```

```

ORIGIN
Query Match      15.1%; Score 250.4; DB 29; Length 844;
Best Local Similarity 77.8%; Pred. No. 2.2e-45;
Matches 343; Conservative 0; Mismatches 88; Indels 10; Gaps 3;

QY 768 TCGGCGATTTCAATCAGTAACATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827
Db 203 TTGGCGACTTCATGTCAGCAACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
QY 828 AATCTGCTAGGATTTGCTGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
Db 263 ATAATCAGCTAAGAATGATCATCTTGAATATGAAAAGGAAAACAGGATTCATAAT 322
QY 886 -----TATTACATAAATATGATGTTCAACATCTTTCTTAATACCTTTTGTCTGATCTCC 939
Db 323 ATCATGTATATAGAGAACATAGTTCAAAAGATCTTCTCAATACCTTTTGTGATCTTCT 382
QY 940 TTAGTAGAGCCAGCGCTGATAGCTGTTAGAACCCCTTGTACCGAGTGTAGTCCAA 999
Db 383 TGAATAGAGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441
QY 1000 -GGCTGCTTAACTGAATAATAGTAGAACACAAGGAGACAAGAGTGTAGAGGAACTGAT 1058
Db 442 GGGCTGCTTAACTGAATAATAGTAGAACACAAGGAGACAAGAGTGTAGAGGAACTGAT 501
QY 1059 TCTTTGTTACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
Db 502 TCTTTGTTACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
QY 1119 TTTATAGACAAACTAGGCTTTTACGCAATATGGCCATATGGCCATATGGCCCAA 1178
Db 562 TTTATAGACAAACTAGGCTTTTACGCAATATGGCCATATGGCCATATGGCCCAA 621
QY 1179 AGGTTTCTTAACTTACCATC 1199
Db 622 AGGTTTCTTAACTTACCATC 642

```

```

RESULT 15
LOCUS CG124360/c
DEFINITION PUFUD11TB ZM_0.6.1.0 KB Zea mays genomic clone ZM5Bc0654A22,
genomic survey sequence.
ACCESSION CG124360
VERSION CG124360.1 GI:34007797
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```


Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 914)
REFERENCE
AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUUFD11TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..914
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0654A22"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match 15.1%; Score 250.2; DB 29; Length 914;
Best Local Similarity 77.8%; Pred. No. 2.4e-45;
Matches 343; Conservative 0; Mismatches 88; Indels 10; Gaps 3;
QY 768 TCGCGGATTTCAATCAGTAACTACATGTCATGCCATCCCATCGACTTCGTGCGCGAT 827
Db 533 TTGGCGACTTCATCTCAGCAACACTCAGCGTTTCATTTCTATCGACTTCGAGTCGGCGAT 474
QY 828 AATCTCCTAGGGATTTGTGTTTTTCATGTATGTGCTTGAATAGAGAGGAAGATAA-- 885
Db 473 ATAATCAGCTAAGAAATGATCATCTTGTATATGAAAAAGGAAAAACAGGATTCAAAAT 414
QY 886 -----TATTACATAAAATATGAGTTCAACATCTTTCTAATACCTTTTGTCTCGATCTCC 939
Db 413 ATCATGTATATAGAGAACTAGTTCAAGATTCTTCTCATATACCTTTGTTGATCTTCT 354
QY 940 TTAGTAGAGCCAGCTGCGTGATACAGTGTAAAGAACCCCTTGTACCGAGTGTAGTCCAA 999
Db 353 TGAATAGAGTCTCATGC-TGATAACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCCAA 295
QY 1000 -GGCTGCTTAACGTAATAATAGTAGAACAACAGGAGACAGAGTGTAGAGAGGAAGTAT 1058
Db 294 GGGCTGCTTAACGTAATAATAGTAGAACAACAGGAGACAGAGTGTAGAGAGGAAGTAT 235
QY 1059 TCCTTTGTTACTATATGGTGTGCTCTCCAAAGGTTACATGATATGGGATCTCCTCTCTA 1118
Db 234 TCCTTTGTTACTATATGGTGTGCTCTCTCCAAAGGTTACATGATATGGGATCTCCTCTCTA 175
QY 1119 TTTATAGACAAAATAGGGTTTCAGGCATATGGGCCATATAGGCCCTTCCTGGCCCAAGAA 1178
Db 174 TTTATAGACAAAATAGGGTTTCAGACATATGGGCCATATGGGCTTCTCTGGCCCAAGAA 115
QY 1179 AGGTTTCTTAACACTACATC 1199
Db 114 AGGTTTCTTAACATAGCAAC 94

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 12:09:10 ; Search time 19145 Seconds
(without alignments)
3753.604 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggtcacaatac.....ctcgttagtttgggacggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	1.5	262181	2	AC125659	AC125659 Rattus no
2	25	1.5	277282	2	AC137423	AC137423 Rattus no
3	23	1.4	173222	9	AC009127	AC009127 Homo sapi
4	23	1.4	174933	9	AC009087	AC009087 Homo sapi
5	23	1.4	184391	2	AC115277	AC115277 Rattus no
6	22	1.3	592	4	AY285636	AY285636 Sus scrofa
7	22	1.3	2409	6	AX653883	AX653883 Sequence
8	22	1.3	109964	2	AP005677	AP005677 Oryza sat
9	22	1.3	150876	2	AC108761	AC108761 Oryza sat
10	22	1.3	151468	2	AC116118	AC116118 Mus muscu
11	22	1.3	176045	10	AC122370	AC122370 Mus muscu
12	22	1.3	191735	10	AL596446	AL596446 Mouse DNA
13	22	1.3	197694	2	AC136065	AC136065 Rattus no
14	22	1.3	197981	2	AC117089	AC117089 Rattus no
15	22	1.3	215867	2	AC099439	AC099439 Rattus no
16	22	1.3	223363	2	AC120700	AC120700 Rattus no
17	22	1.3	242309	2	AC098167	AC098167 Rattus no
18	22	1.3	245893	2	AC125267	AC125267 Mus muscu
19	22	1.3	256105	2	AC093979	AC093979 Rattus no
20	22	1.3	261414	2	AC122960	AC122960 Rattus no
21	22	1.3	277860	2	AC124647	AC124647 Mus muscu
22	22	1.3	290002	2	AC134106	AC134106 Rattus no
23	22	1.3	294415	2	AC110321	AC110321 Rattus no
24	21	1.3	426	8	AF218307	AF218307 Saccharom
25	21	1.3	802	8	AY089172	AY089172 Arabidops
26	21	1.3	3717	1	D49692	D49692 Spirulina p
27	21	1.3	5956	6	AX251973	AX251973 Sequence
28	21	1.3	5956	6	AX344365	AX344365 Sequence
29	21	1.3	5956	6	AX348758	AX348758 Sequence
30	21	1.3	11915	1	AE009922	AE009922 Pyrobacul
31	21	1.3	28573	3	AF024502	AF024502 Caenorhab
32	21	1.3	31219	2	AC019857	AC019857 Drosophil
33	21	1.3	39000	9	AP003024	AP003024 Homo sapi
34	21	1.3	69646	2	AC102089	AC102089 Mus muscu
35	21	1.3	73746	2	AC101286	AC101286 Mus muscu
36	21	1.3	93216	9	AC003016	AC003016 Human BAC
37	21	1.3	98535	9	AF188025	AF188025 Homo sapi
38	21	1.3	99150	9	AF202964	AF202964 Homo sapi
39	21	1.3	112070	9	AC068798	AC068798 Homo sapi
40	21	1.3	119654	2	AC134325	AC134325 Mus muscu
41	21	1.3	129779	3	AC006214	AC006214 Drosophil
42	21	1.3	134784	8	FL1123	FL1123 Arabidops
43	21	1.3	143441	2	AC013369	AC013369 Homo sapi
44	21	1.3	144290	9	AC034251	AC034251 Homo sapi
45	21	1.3	146811	9	AF121898	AF121898 Homo sapi

ALIGNMENTS

RESULT 1
AC125659
LOCUS
AC125659
DEFINITION
Rattus norvegicus clone CH230-11B1, WORKING DRAFT SEQUENCE.
AC125659
ACCESSION
AC125659.3 GI:24940742
VERSION
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 262181)
Muzny, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deyam, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Guevara, W., Guerrero, C., Giller, R., Grady, M., Hamilton, C., Hamilton, K., Gunaratne, P., Haaland, W., Hamil, C., Henderson, N., Hernandez, J., Harvey, Y., Hawlak, P., Hawes, A., Henderson, N., Hernandez, J., Lorensuwa, L., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwako, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 262181)
Unpublished
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 262181)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:22772668.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GPVZ
Center clone name: CH230-11B1
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 232291 bases at least Q40
Consensus quality: 235081 bases at least Q30
Consensus quality: 236925 bases at least Q20
Estimated insert size: 238350; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 262181: contig of 262181 bp in length.
FEATURES
source
1..262181
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-11B1"
misc_feature
1..1034
/note="wgs_contig"
misc_feature
259169..261036
/note="wgs_contig"
misc_feature
261087..262181
/note="wgs_contig"
ORIGIN
Query Match 1.5% Score 25; DB 2; Length 262181;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1167 CTGGCCCAAGAAAGGTTTCTTACA 1191
|||||
DB 242133 CTGGCCCAAGAAAGGTTTCTTACA 242157
RESULT 2
AC137423
LOCUS
DEFINITION
AC137423 Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
*** 7 unordered pieces.
AC137423
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 27282)
Muzny, D., Marle, Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogis, E., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hernandez, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hoques, M., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Jackson, A., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kowar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, N., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, J., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, I., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Unpublished
Direct Submission
2 (bases 1 to 277282)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

Center project name: KZLE
Center clone name: CH230-unknown
Assembly program: Phrap; version 0.990329
Consensus quality: 204050 bases at least Q40

Consensus quality: 209444 bases at least Q30
Consensus quality: 212955 bases at least Q20
Estimated insert size: 204859; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 12348: contig of 12348 bp in length
* 12349 12448: gap of unknown length
* 12449 251860: contig of 239412 bp in length
* 251861 251960: gap of unknown length
* 251961 255729: contig of 3769 bp in length
* 255730 255829: gap of unknown length
* 255830 268931: contig of 13102 bp in length
* 268932 269031: gap of unknown length
* 269032 27015: contig of 1684 bp in length
* 27016 270815: gap of unknown length
* 270816 274240: contig of 3425 bp in length
* 274241 274340: gap of unknown length
* 274341 277282: contig of 2942 bp in length.

FEATURES
source

1. 277282
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-unknown"
42559. 44206
/note="wgs_contig"
97204. .99317
/note="wgs_contig"
115739. .118162
/note="wgs_contig"
255830. 257305
/note="wgs_contig"
268832. .268931
/note="wgs_contig"

ORIGIN

Query Match 1.5% Score 25; DB 2; Length 277282;
Best Local Similarity 100.0%; Pred.No. 0.17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1167 CTGCCCCCAAGAAAGGTTTCTTAACA 1191
|||||
Db 84796 CTGCCCCCAAGAAAGGTTTCTTAACA 84820

RESULT 3
AC009127/c

LOCUS AC009127 173222 bp DNA linear PRI 12-APR-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-498D10, complete sequence.
ACCESSION AC009127
VERSION AC009127.9 GI:20143529
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
1 (bases 1 to 173222)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 173222)
DOE Joint Genome Institute.
Direct Submission

JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 173222)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Apr 12, 2002 this sequence version replaced gi:13786311.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

source

1..173222

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-498D10"

ORIGIN

Query Match 1.4%; Score 23; DB 9; Length 173222;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1546 CAAAGCAGCAGCACAGCCCAACA 1568

|||||

Db 45290 CAAAGCAGCAGCACAGCCCAACA 45268

RESULT 4

AC009087/c

LOCUS AC009087 174933 bp DNA linear PRI 27-APR-2000

DEFINITION Homo sapiens chromosome 16 clone RP11-384M15, complete sequence.

ACCESSION AC009087

VERSION AC009087.4 GI:7656694

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 174933)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 174933)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 174933)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Submitted (27-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Apr 27, 2000 this sequence version replaced gi:7025737.

Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center

www.shgc.stanford.edu

Quality: Phrap Quality >=40 99.7% of Sequence;

Estimated Total Number of Errors is 1.

STS Content:

SHGC-32168 G27137

WI-15979 G21275

WI-10173 G11750

WI-17126 G21504

WI-9359 G06054.

Location/Qualifiers

1..174933

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-384M15"

ORIGIN

Query Match 1.4%; Score 23; DB 9; Length 174933;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1546 CAAAGCAGCAGCACAGCCCAACA 1568

|||||

Db 150280 CAAAGCAGCAGCACAGCCCAACA 150258

RESULT 5

AC115277/c

LOCUS AC115277 184391 bp DNA linear HTG 15-NOV-2002

DEFINITION Rattus norvegicus clone CH230-372C1, *** SEQUENCING IN PROGRESS ***.

ACCESSION AC115277

VERSION AC115277.5 GI:25009102

KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT; HTGS_ENRICHED.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (bases 1 to 184391)

Muzny, D., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A., Chu, J., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Fan, G., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Grabegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., Hodge, A., Hodge, A., Hodge, A., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhsuwa, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, P., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedemeh, O., Okwunonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, J., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, K., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 184391)
Worley, K.C.

Direct Submission
Submitted (16-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 184391)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23194609.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequences contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTHP
Center clone name: CH230-372C1
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 16483 bases at least Q40
Consensus quality: 167335 bases at least Q30
Consensus quality: 169474 bases at least Q20
Estimated insert size: 170619; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*
FEATURES
source
1 184391: contig of 184391 bp in length.
Location/Qualifiers
1. 184391
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-372C1"
1. 1601
/note="wgs end extension
clone end:Sp6"
1705_3257
/note="wgs end extension
clone end:Sp6"
4527..107311
/note="clone boundary
clone end:Sp6"
site:
end sequence:BZ216963"
170821..173036
/note="wgs contig"
complement(182070..182951)
/note="clone boundary
clone end:T7"
site:
end sequence:BZ216960"
ORIGIN
Query Match 1.4%; Score 23; DB 2; Length 184391;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 842 TTTGTGCTTTTCATGTATGTGCT 864
|||||
Db 77557 TTTGTGCTTTTCATGTATGTGCT 77535
|||||
RESULT 6
AY285636/c
LOCUS AY285636 592 bp DNA linear MAM 30-JUN-2003
DEFINITION Sus scrofa clone UMNp983.gcg microsatellite sequence.
ACCESSION AY285636
VERSION AY285636.1 GI:32347979
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Alexander, L.J., Morrison, L.Y. and Farenkrug, S.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2003) Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave, St. Paul, MN 55108, USA
FEATURES
source
1. 592
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="UMNp983.gcg"
1. 592
repeat_region
/note="microsatellite"
/rpt_type=tandem
ORIGIN
Query Match 1.3%; Score 22; DB 4; Length 592;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 98 GAAAGCAAAATATTTTAAATGAA 119
|||||
Db 546 GAAAGCAAAATATTTTAAATGAA 525
|||||
RESULT 7

```

AX553883/c
LOCUS      AX553883      2409 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 3753 from Patent WO03000898.
ACCESSION  AX553883
VERSION     AX553883.1  GI:29156697
SOURCE      Oryza sativa
ORGANISM    Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
AUTHORS     Chang H.S., Chen W., Cooper B., Glazebrook J., Goff S.A., Hou Y.M.,
            Katagiri F., Qian S., Tao Y., Whitam S., Xie Z., Zhu T. and Zou G.
TITLE       Plant genes involved in defense against pathogens
JOURNAL     Patent: WO 03000898-A 3753 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES   Location/Qualifiers
            1..2409
            /organism="Oryza sativa"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4530"
ORIGIN
Query Match      1..3%; Score 22; DB 6; Length 2409;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      657 CGACGACGTCGACGCTGTGCT 678
        |||||
Db      492 CGACGACGTCGACGCTGTGCT 471

RESULT 8
AP005677/c
LOCUS      AP005677      109964 bp      DNA      linear      HTG 06-SEP-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
            OJ1011 C06, *** SEQUENCING IN PROGRESS ***
ACCESSION  AP005677
VERSION     AP005677.1  GI:22775435
            HTG; HTGS, PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
AUTHORS     Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE       Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
            clone: OJ1011 C06
JOURNAL     Published Only in Database (2002)
REFERENCE   2 (bases 1 to 109964)
AUTHORS     Sasaki, T., Matsumoto, T. and Katayose, Y.
TITLE       Direct Submission
JOURNAL     Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
            Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            1..109964
FEATURES   Location/Qualifiers
source
1..109964

AX553883/c
LOCUS      AX553883      2409 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 3753 from Patent WO03000898.
ACCESSION  AX553883
VERSION     AX553883.1  GI:29156697
SOURCE      Oryza sativa
ORGANISM    Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
AUTHORS     Chang H.S., Chen W., Cooper B., Glazebrook J., Goff S.A., Hou Y.M.,
            Katagiri F., Qian S., Tao Y., Whitam S., Xie Z., Zhu T. and Zou G.
TITLE       Plant genes involved in defense against pathogens
JOURNAL     Patent: WO 03000898-A 3753 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES   Location/Qualifiers
            1..2409
            /organism="Oryza sativa"
            /mol_type="unassigned DNA"
            /db_xref="taxon:4530"
ORIGIN
Query Match      1.3%; Score 22; DB 2; Length 109964;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      657 CGACGACGTCGACGCTGTGCT 678
        |||||
Db      77096 CGACGACGTCGACGCTGTGCT 77075

RESULT 9
AC108761/c
LOCUS      AC108761      150876 bp      DNA      linear      HTG 31-JAN-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
            OSUNBa0087J09, *** SEQUENCING IN PROGRESS ***; 7 unordered pieces.
ACCESSION  AC108761
VERSION     AC108761.1  GI:18449966
            HTG; HTGS, PHASE1.
SOURCE      Oryza sativa (japonica cultivar-group)
ORGANISM    Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 150876)
AUTHORS     Lee, J.-S., Hahn, J.-H., Yoon, U.-H., Yun, D.-W., Lee, M.-C., Eun, M.-Y.
            and Kim, H.-I.
TITLE       Oryza sativa BAC OSUNBa0087J09 genomic sequence
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 150876)
AUTHORS     Hahn, J.-H. and Kim, H.-I.
TITLE       Direct Submission
JOURNAL     Submitted (31-JAN-2002) Rice Genome Sequencing Project, National
            Institute of Agricultural Science and Technology (NIAS), RDA, 249
            Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@rda.go.kr,
            Tel: 82-31-290-0309, Fax: 82-31-290-0308)
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * 1 45839: contig of 45839 bp in length
            * 45840 45939: gap of unknown length
            * 45940 76833: contig of 30894 bp in length
            * 76834 76933: gap of unknown length
            * 76934 109205: contig of 32272 bp in length
            * 109206 109305: gap of unknown length
            * 109306 132339: contig of 22934 bp in length
            * 132340 132339: gap of unknown length
            * 132340 134166: contig of 1827 bp in length
            * 134167 134266: gap of unknown length
            * 134267 140337: contig of 5971 bp in length
            * 140338 140337: gap of unknown length
            * 140338 150876: contig of 10539 bp in length.
FEATURES   Location/Qualifiers
source
1..150876
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/variety="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="OSUNBa0087J09"
ORIGIN

```

Query Match 1.3%; Score 22; DB 2; Length 150876;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGCTGATGCT 678
 DB 118775 CGACGACGTCGACGCTGATGCT 118754

RESULT 10
 AC116118
 LOCUS
 DEFINITION Mus musculus chromosome 16 clone RP23-73P19 map 16, *** SEQUENCING
 IN PROGRESS ***, 8 unordered pieces.
 AC116118
 AC116118.5 GI:39841503
 HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 151468)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome 16, clone RP23-73P19
 Unpublished
 2 (bases 1 to 151468)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,A., Collangelo,M., Collins,S., Collamore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,M., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kellis,C., Larocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., MacDonald,P., Major,J., Marcuis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menesius,L.,
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 151468)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collamore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K., Faro,S.,
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faros,S.,
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
 MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrim,J., Menesius,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission

TITLE

JOURNAL
 COMMENT
 Submitted (15-DEC-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Dec 15, 2003 this sequence version replaced gi:28460918.
 All repeats were identified using RepeatMasker.
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: h23300
 Center clone name: 73_P_19

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 13387: contig of 13387 bp in length
 * 13388 13487: gap of 100 bp
 * 13488 18518: contig of 5031 bp in length
 * 18519 18618: gap of 100 bp
 * 18619 35043: contig of 16425 bp in length
 * 35044 35143: gap of 100 bp
 * 35144 54747: contig of 19604 bp in length
 * 54748 54847: gap of 100 bp
 * 54848 64657: contig of 9910 bp in length
 * 64658 64758: gap of 100 bp
 * 64758 101040: contig of 36283 bp in length
 * 101041 101140: gap of 100 bp
 * 101141 130194: contig of 29054 bp in length
 * 130195 130294: gap of 100 bp
 * 130295 151468: contig of 21174 bp in length.

FEATURES

Location/Qualifiers
 1. 151468
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="16"
 /map="16"
 /clone_lib="RP23-73P19"
 /clone="RP23-73P19" Female Mouse BAC"

ORIGIN

Query Match 1.3%; Score 22; DB 2; Length 151468;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 TTTTGTCTCTCTCTCTCTCTCTG 711
 |||||

Db 51549 TTTTGTCTCTCTCTCTCTCTG 51570
 |||||

RESULT 11

AC122370
 LOCUS
 DEFINITION Mus musculus BAC clone RP23-45988 from chromosome 8, complete
 sequence.
 AC122370
 AC122370.4 GI:28376820
 HTG.
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 176045)
 Berghoff,A., Haakenson,W., Bielicki,L. and Meyer,R.
 The sequence of Mus musculus BAC clone RP23-45988

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (2001)
2 (bases 1 to 176045)
Wilson,R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 176045)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 176045)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (29-JAN-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 176045)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (14-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 176045)
Wilson,R.
Direct Submission
Submitted (11-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 14, 2003 this sequence version replaced gi:28016323.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@watson.wustl.edu
----- Summary Statistics

Center project name: M_BA0459B08

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
The RP23-23 BAC Library has been constructed by Kazutoyo Oseegawa
and Minako Tateno in the laboratory of Pieter de Jong
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or
brain genomic DNA. The clone and detailed information can be
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC117189.

FEATURES
source
1. .176045
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone="RP23-459B8"

repeat_region
1..165
/clone_lib="RPCI-23"
/rpt_family="MaLR"
259..397
/rpt_family="L1"
398..1102
/rpt_family="L1"
1146..1270
/rpt_family="L1"
1352..1542
/rpt_family="L1"
1549..1661
/rpt_family="L1"
1668..1753
/rpt_family="L1"
1819..3363
/rpt_family="L1"
4365..4836
/rpt_family="MER2_type"
7124..7496
/rpt_family="ERV1"
7598..8149
/rpt_family="L1"
8553..8974
/rpt_family="L1"
8969..9414
/rpt_family="L1"
9411..9561
/rpt_family="L1"
9722..10175
/rpt_family="L1"
10233..10528
/rpt_family="ERVK"
10529..11178
/rpt_family="L1"
11900..12045
/rpt_family="L1"
12498..12616
/rpt_family="ERV1"
13172..19405
/rpt_family="MaLR"
21902..22009
/rpt_family="Alu"
22010..22069
/rpt_family="MaLR"
23313..23416
/rpt_family="Alu"
23970..24770
/rpt_family="L1"
25850..26050
/rpt_family="B2"
26146..26323
/rpt_family="B2"
27097..27228
/rpt_family="L1"
27582..28676
/rpt_family="L1"
28684..28861
/rpt_family="B2"
28892..29358
/rpt_family="L1"
29440..29577
/rpt_family="Alu"
29608..29744
/rpt_family="Alu"
29761..30179
/rpt_family="L1"
30176..30307
/rpt_family="L1"
30298..30453
/rpt_family="L1"
30454..30801
/rpt_family="ERVK"


```

repeat_region 30802..31095
/rpt_family="L1"
repeat_region 31104..31287
/rpt_family="L1"
repeat_region 31663..33668
/rpt_family="L1"
repeat_region 33667..33893
/rpt_family="L1"
repeat_region 33893..34068
/rpt_family="L1"
repeat_region 34072..34262
/rpt_family="B2"
repeat_region 34282..34355
/rpt_family="L1"
repeat_region 35311..35384
/rpt_family="MaLR"
repeat_region 35526..35938
/rpt_family="MaLR"
repeat_region 35976..36243
/rpt_family="L1"
repeat_region 36413..37025
/rpt_family="L1"
repeat_region 37081..37528
/rpt_family="L1"
repeat_region 37668..37812
/rpt_family="L1"
repeat_region 37893..38081
/rpt_family="Alu"
repeat_region 38203..38420
/rpt_family="B4"
repeat_region 39026..39165
/rpt_family="MIR"
repeat_region 40708..40898
/rpt_family="MaLR"
repeat_region 40947..41307
/rpt_family="L1"
repeat_region 41599..41691
/rpt_family="Alu"
repeat_region 43021..43288
/rpt_family="L1"
repeat_region 43259..43417

```

Query Match 1.3%; Score 22; DB 10; Length 176045;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TATTTTCTTAAAGAAAAA 248
 Db 88207 TATTTTCTTAAAGAAAAA 88228

RESULT 12
 AC1596446
 LOCUS AC1596446 191735 bp DNA linear ROD 05-APR-2002
 DEFINITION Mouse DNA sequence from clone RP23-386E10 on chromosome 11,
 complete sequence.
 ACCESSION AL596446
 VERSION AL596446.11 GI:20068511
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Tracey, A.
 Direct Submission
 Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:17043843.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em; EMBL; SW;
 SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-386E10 is
 from the RPCI-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6.

FEATURES

Location/Qualifiers
 1..191735
 /organism="Mus musculus"
 /mol_type="Genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-386E10"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 1.3%; Score 22; DB 10; Length 191735;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 TTTTGTTCCTTCTCTCTG 711
 Db 157107 TTTTGTTCCTTCTCTG 157128

RESULT 13
 AC136065/c

LOCUS AC136065 197694 bp DNA linear HTG 29-OCT-2002
 DEFINITION Rattus norvegicus clone CH230-85M19, *** SEQUENCING IN PROGRESS
 ***, 85 unordered pieces.

ACCESSION AC136065.1 GI:24417936

VERSION HTG; HTGS PHASE1.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

1 (bases 1 to 197694)

REFERENCE

AUTHORS

Muzny, D., Maric, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Bialwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.W., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gubregoeis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevata, W.,
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

TITLE	REFERENCE	AUTHORS	JOURNAL
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,	8891	contig of 1097 bp in length	
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,	9988	10087: gap of unknown length	
Li, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,	10088	11435: contig of 1348 bp in length	
Lorensuhwa, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J.,	11436	11535: gap of unknown length	
Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A.,	11536	12592: contig of 1057 bp in length	
Mangum, B., Mapus, P., Martin, K., Martin, R., Martinez, E.,	12593	12692: gap of unknown length	
Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,	12693	13986: contig of 1304 bp in length	
Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,	13987	14096: gap of unknown length	
Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,	14097	15658: gap of unknown length	
Morris, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G.,	15659	15759: contig of 1203 bp in length	
Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,	15759	16961: gap of unknown length	
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,	16962	17061: contig of 1212 bp in length	
Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E.,	17062	18273: gap of unknown length	
Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,	18274	19985: contig of 1612 bp in length	
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,	19986	20085: gap of unknown length	
Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S.,	20086	21564: contig of 1479 bp in length	
Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,	21565	22627: contig of 1263 bp in length	
Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E.,	22628	23027: gap of unknown length	
Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,	23028	24683: contig of 1656 bp in length	
Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,	24684	24783: gap of unknown length	
Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,	24784	26440: contig of 1657 bp in length	
Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,	26441	26540: gap of unknown length	
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,	26541	28249: contig of 1709 bp in length	
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,	28250	28349: gap of unknown length	
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,	28350	29394: contig of 1045 bp in length	
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R.,	29395	31413: contig of 1919 bp in length	
Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.	31414	31513: gap of unknown length	
Direct Submission	31514	34099: contig of 2586 bp in length	
Unpublished	34100	34199: gap of unknown length	
2 (bases 1 to 197694)	34200	36752: contig of 2553 bp in length	
Rat Genome Sequencing Consortium.	36753	38733: contig of 1521 bp in length	
Direct Submission	38734	40434: contig of 1961 bp in length	
Submitted (29-OCT-2002) Human Genome Sequencing Center, Department	40435	40534: gap of unknown length	
of Molecular and Human Genetics, Baylor College of Medicine, One	40535	41942: contig of 1408 bp in length	
Baylor Plaza, Houston, TX 77030, USA	41943	42042: gap of unknown length	
----- Genome Center	42043	43635: contig of 1593 bp in length	
Center: Baylor College of Medicine	43636	43735: gap of unknown length	
Center code: BCM	43736	45768: contig of 2033 bp in length	
Web site: http://www.hgsc.bcm.tmc.edu/	45769	45868: gap of unknown length	
Contact: hgsc-help@bcm.tmc.edu	45869	47681: contig of 1813 bp in length	
----- Project Information	47682	47781: gap of unknown length	
Center project name: KDBI	47782	50178: contig of 2397 bp in length	
Center clone name: CH230-85M19	50179	50278: gap of unknown length	
----- Summary Statistics	50279	53070: contig of 2792 bp in length	
Sequencing vector: Plasmid	53071	53170: gap of unknown length	
Chemistry: Dye-terminator Big Dye: 100% of reads	53171	56532: contig of 3362 bp in length	
Assembly program: Phrap: version 0.990329	56533	56632: gap of unknown length	
Consensus quality: 146097 bases at least Q40	56633	58421: contig of 1789 bp in length	
Consensus quality: 151106 bases at least Q30	58422	58521: gap of unknown length	
Consensus quality: 155140 bases at least Q20	58522	60418: contig of 1897 bp in length	
-----	60419	60518: gap of unknown length	
* NOTE: Estimated insert size may differ from sequence length	60519	62932: contig of 2414 bp in length	
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).	62933	63032: gap of unknown length	
* NOTE: This is a 'working draft' sequence. It currently	63033	65252: contig of 2220 bp in length	
* consists of 65 contigs. The true order of the pieces	65253	65352: gap of unknown length	
* is not known and their order in this sequence record is	65353	67620: contig of 2268 bp in length	
* arbitrary. Gaps between the contigs are represented as	67621	67720: gap of unknown length	
* runs of N, but the exact sizes of the gaps are unknown.	67721	69946: contig of 2226 bp in length	
* This record will be updated with the finished sequence	69947	70046: gap of unknown length	
* as soon as it is available and the accession number will	70047	73197: contig of 3151 bp in length	
* be preserved.	73198	73297: gap of unknown length	
1 1241: contig of 1241 bp in length	73298	75571: contig of 2274 bp in length	
1242 1341: gap of unknown length	75572	75672: gap of unknown length	
1342 2851: contig of 1510 bp in length	75672	78398: contig of 2727 bp in length	
2852 2951: gap of unknown length	78399	78498: gap of unknown length	
2952 4172: contig of 1221 bp in length	7849		

```

* 84947 85046: gap of unknown length
* 85047 88159: contig of 3113 bp in length
* 88160 88259: gap of unknown length
* 88260 92104: contig of 3845 bp in length
* 92105 92204: gap of unknown length
* 92205 96596: contig of 4392 bp in length
* 96597 96596: gap of unknown length
* 96597 100965: contig of 4269 bp in length
* 100966 101065: gap of unknown length
* 101066 104609: contig of 3544 bp in length
* 104610 109250: contig of 4541 bp in length
* 109251 109350: gap of unknown length
* 109351 113836: contig of 4486 bp in length
* 113837 113936: gap of unknown length
* 113937 118684: contig of 4648 bp in length
* 118685 118684: gap of unknown length
* 118685 122698: contig of 4014 bp in length
* 122699 122798: gap of unknown length
* 122799 127266: contig of 4468 bp in length
* 127267 127366: gap of unknown length

Query Match 1.3%; Score 22; DB 2; Length 197694;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 TCTATTTTCTTAAAGAAAAA 246
Db 40901 TCTATTTTCTTAAAGAAAAA 40880

RESULT 14
AC117089
LOCUS
DEFINITION
AC117089 197981 bp DNA linear HTG 15-NOV-2002
Rattus norvegicus clone CH230-11969, WORKING DRAFT SEQUENCE, 2
unordered pieces.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 197981)
Muzny D.M., Metzker M.L., Abrahams S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Delgado O., Denson S., Detamo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregeorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,
Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hognes M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird P., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpathy S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorenshewa L., Loulsegad H., Lozano R.J., Lu X., Ma J.,
Maheshwari M., Mahindran M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapa P., Martin K., Martin R., Martinez E.,
Mawhiney S., McLeod M.P., McNeill T.Z., Meenen E.,

```

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarpunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
REFERENCE
AUTHORS
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 197981)
Worley, K.C.
Direct Submission
Submitted (08-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 197981)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23194925.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GFPD
Center clone name: CH230-11969
Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 189549 bases at least Q40
Consensus quality: 190935 bases at least Q30
Consensus quality: 192081 bases at least Q20
Estimated insert size: 196342; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
 * 1 196422: contig of 196422 bp in length
 * 196423 196522: gap of unknown length
 * 196523 197981: contig of 1459 bp in length.

FEATURES

Source

1. 197981
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-119G9"
 194929. 195819
 /note="clone boundary
 clone_end:T7
 site:
 end_sequence:BH315129"

misc_feature

ORIGIN

Query Match 1.3%; Score 22; DB 2; Length 197981;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 TTTTGTGTTTGTCTCTGTG 711

Db 53534 TTTTGTGTTTGTCTCTGTG 53555
 |||||
 |||||

RESULT 15

AC099439/c

LOCUS

Rattus norvegicus clone CH230-145P16, WORKING DRAFT SEQUENCE. HTG 10-MAY-2003

DEFINITION

AC099439

AC099439.5 GI:30521297

HTG; HTGS_PHASE2; HTGS DRAFT; HTGS_FULLTOP.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 215867)
 Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Cartroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howells,S., Huliy,S., Hume,J., Idiebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhewa,L., Louised,H., Lozado,R.J., Lu,X., Ma,J.,
 Mangum,B., Mapua,P., Martin,K., Mahmoud,M., Malloy,K., Mangum,A.,
 Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Martine,E.,
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwakoeleneh,O., Okwuono,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
 Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Submitted (15-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GIHG
 Center clone name: CH230-145P16
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 210812 bases at least Q40
 Consensus quality: 212267 bases at least Q30
 Consensus quality: 213291 bases at least Q20
 Estimated insert size: 221076; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 215867: contig of 215867 bp in length.

FEATURES

source

Location/Qualifiers
 1. 215867
 /organism="Rattus norvegicus"

```

/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-145P16"
1..1660
/note="wgs_contig"

```

misc_feature

ORIGIN

```

Query Match      1.3%; Score 22; DB 2; Length 215867;
Best Local Similarity 100.0%; Pred.No. 7.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      690 TTTTGTGTTTGTGTTTCTCTGTG 711
          |||||||
Db      46567 TTTTGTGTTTGTGTTTCTCTGTG 46546

```

Search completed: September 2, 2004, 20:00:48
Job time : 19151 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 10:43:49 ; Search time 2015 Seconds
(without alignments)
3495.540 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattacggctcacatac.....ctcgttagttggagcggcg 1658

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1980s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	1.3	2409	7	ADA70430 Rice gene
C 2	21	1.3	801	3	AAC50800 Arabidops
C 3	21	1.3	802	3	AAC33054 Arabidops
C 4	21	1.3	5956	6	ABK31369 Signal tr
C 5	21	1.3	5956	6	ABL70326 Chemical
C 6	21	1.3	5956	6	AAS61273 Human gen
C 7	21	1.3	16739	4	ABL07848 Drosophil
C 8	21	1.3	21828	4	ABL07850 Drosophil
C 9	20	1.2	291	6	ABN17338 Human ORF
C 10	20	1.2	785	4	AAL15883 Human bre
C 11	20	1.2	830	8	ADA49079 Wheat gen
C 12	20	1.2	1044	7	ACC61661 Gene sequ
C 13	20	1.2	1963	3	AAZ52423 HTRM clon
C 14	20	1.2	1973	4	AAL60732 Human pol
C 15	20	1.2	1979	4	AAL58946 Human pol
C 16	20	1.2	1979	8	ADH48928 Novel hum
C 17	20	1.2	3255	7	ADA53534 Human cod
C 18	20	1.2	4633	4	AAL93684 Human TNF
C 19	20	1.2	5823	6	ABK28382 DNA trans
C 20	20	1.2	12988	4	AAK80680 Human imm
C 21	20	1.2	12988	7	ABZ74054 Secreted
C 22	20	1.2	12988	9	ADC20775 Human sec
C 23	20	1.2	16163	4	AAL62620 Human bre

24	20	1.2	16163	4	AAK83901	Aak83901 Human imm
25	20	1.2	16163	4	AAK87161	Aak87161 Human imm
c 26	20	1.2	16163	4	AAL03817	Aal03817 Human rep
27	20	1.2	100301	6	ABQ88176	Abq88176 Human ost
c 28	19	1.1	558	3	AAF08066	Aaf08066 Fusarium
c 29	19	1.1	604	6	ABN87803	Abn87803 Human ova
c 30	19	1.1	648	7	ABZ54983	Abz54983 Aspergill
c 31	19	1.1	734	6	ABK85714	Abk85714 DNA encod
c 32	19	1.1	771	4	AAK59627	Aak59627 Human imm
c 33	19	1.1	911	6	ABK99945	Abk99945 DNA encod
c 34	19	1.1	1087	4	AAK75762	Aak75762 Human imm
c 35	19	1.1	1447	2	AAV08179	Aav08179 Bovine in
c 36	19	1.1	1447	2	AAV68162	Aav68162 Bovine in
c 37	19	1.1	1449	1	AAN30041	Aan30041 Sequence
c 38	19	1.1	1449	6	ABS53995	Abs53995 cDNA enco
c 39	19	1.1	1469	1	AAN30039	Aan30039 Sequence
c 40	19	1.1	1469	2	AAV08177	Aav08177 Bovine in
c 41	19	1.1	1469	2	AAV68160	Aav68160 Bovine in
c 42	19	1.1	1469	6	ABS53993	Abs53993 cDNA enco
c 43	19	1.1	1484	9	ADD30329	Add30329 Plant yie
c 44	19	1.1	1551	6	ABZ31860	Abz31860 Candida a
c 45	19	1.1	1557	6	ABQ67825	Abq67825 Listeria

ALIGNMENTS

RESULT 1

ID	ADA70430/c
XX	ADA70430 standard; DNA; 2409 BP.
AC	ADA70430;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Rice gene, SEQ ID 3753.
XX	
KW	Plant; bacterial infection; fungal infection; viral infection; rice;
KW	gene; ds.
XX	
OS	Oryza sativa.
XX	
PN	WO2003000898-A1.
XX	
FD	03-JAN-2003.
PF	
XX	
XX	22-JUN-2001; 2001WO-IB001105.
PR	
XX	22-JUN-2001; 2001WO-IB001105.
FA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX	
DR	WPI; 2003-175296/17.
XX	
PT	Identifying at least one gene involved in plant resistance or response to
PT	pathogenic infection for conferring resistance or tolerance to a plant to
PT	bacterial, fungal or viral infection by determining or detecting plant
PT	gene expression.
XX	
PS	Claim 6; SEQ ID NO 3753; 899pp; English.
XX	

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to

CC illustrate the invention.

XX Sequence 2409 BP; 399 A; 846 C; 799 G; 361 T; 0 U; 4 Other;

Query Match 1.3%; Score 22; DB 7; Length 2409;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTCGACGTCGTCGTC 678

Db 492 CGACGACGTCGACGTCGTCGTC 471

RESULT 2

ID AAC50800/c

AC AAC50800;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 66182.

DE Hybridisation assay; Genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 10-JUN-1999; 99US-0138847P.

PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140821P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

Qy 691 TTGTTTTTCTCTG 711
 Db 24 TTGTTTTTCTCTG 4

RESULT 3
 AAC33054/c
 ID AAC33054 standard; DNA; 802 BP.
 XX AC AAC33054;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 1614.
 XX KW Hybridisation assay; genetic mapping; gene expression control;
 XX KW protein identification; signal transduction pathway; metabolic pathway;
 XX KW promoter; termination sequence; ss.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-00301439.
 XX PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 29-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.

Query Match 1.3%; Score 21; DB 3; Length 801;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-014287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0145386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149823P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155653P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158023P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159323P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 1.3%; Score 21; DB 3; Length 802;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 691 TTGTTTGTCTCTGCGTG 711
 Db 24 TTGTTTGTCTCTGCGTG 4

ABK31369
 ID ABK31369 standard; DNA; 5956 BP.
 XX
 AC ABK31369;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Signal transduction associated gene modified complementary DNA #106.
 XX
 KW Human; signal transduction associated gene; cytosine methylation state;
 KW CpG island; signal transduction associated disease; solid tumour; cancer;
 KW antitumour; cytostatic; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 XX WO200200926-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX
 XX 29-JUN-2001; 2001WO-EP007472.
 PF
 XX 30-JUN-2000; 2000DE-01032529.
 PR
 XX 01-SEP-2000; 2000DE-01043826.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 FI
 XX WPI; 2002-147896/19.
 DR
 XX
 XX Oligonucleotide for diagnosis and therapy of diseases associated with
 PT signal transduction e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with signal transduction.
 XX
 PS Claim 1; SEQ ID NO 212; 24pp; English.
 XX
 CC The present invention relates to chemically modified DNA sequences of
 CC signal transduction associated genes. The DNA sequences are chemically
 CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
 CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
 CC the cytosine methylation state (CpG islands) of these genes, and a method
 CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
 CC genes associated with signal transduction. The genomic DNA can be
 CC obtained from cells or cellular components which contain DNA, e.g. cell
 CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
 CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
 CC brain, heart, prostate, lung, breast or liver, histologic object slides,
 CC and all their possible combinations. The sequences of the invention are
 CC useful for the diagnosis and therapy of diseases associated with signal
 CC transduction e.g. solid tumours and cancer. ABK31369-ABK31374 represent
 CC chemically pretreated genomic DNA sequences of different genes associated
 CC with signal transduction, or their complementary sequences. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office
 XX
 SQ Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 6; Length 5956;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 287 TTTTATATATATAGATAT 307
 Db 5381 TTTTATATATATAGATAT 5401
 RESULT 5
 ABL70326
 ID ABL70326 standard; DNA; 5956 BP.
 XX
 AC ABL70326;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Signal transduction associated gene modified complementary DNA #106.
 XX
 KW Human; signal transduction associated gene; cytosine methylation state;
 KW CpG island; signal transduction associated disease; solid tumour; cancer;
 KW antitumour; cytostatic; mutant; ds.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 XX WO200200926-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX
 XX 29-JUN-2001; 2001WO-EP007472.
 PF
 XX 30-JUN-2000; 2000DE-01032529.
 PR
 XX 01-SEP-2000; 2000DE-01043826.
 PR
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 FI
 XX WPI; 2002-147896/19.
 DR
 XX
 XX Oligonucleotide for diagnosis and therapy of diseases associated with
 PT signal transduction e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with signal transduction.
 XX
 PS Claim 1; SEQ ID NO 212; 24pp; English.
 XX
 CC The present invention relates to chemically modified DNA sequences of
 CC signal transduction associated genes. The DNA sequences are chemically
 CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.
 CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting
 CC the cytosine methylation state (CpG islands) of these genes, and a method
 CC for the diagnosis and/or therapy of genetic and epigenetic parameters of
 CC genes associated with signal transduction. The genomic DNA can be
 CC obtained from cells or cellular components which contain DNA, e.g. cell
 CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,
 CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,
 CC brain, heart, prostate, lung, breast or liver, histologic object slides,
 CC and all their possible combinations. The sequences of the invention are
 CC useful for the diagnosis and therapy of diseases associated with signal
 CC transduction e.g. solid tumours and cancer. ABK31369-ABK31374 represent
 CC chemically pretreated genomic DNA sequences of different genes associated
 CC with signal transduction, or their complementary sequences. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office
 XX
 SQ Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 6; Length 5956;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 287 TTTTATATATATAGATAT 307
 Db 5381 TTTTATATATATAGATAT 5401
 RESULT 6
 AAS61273
 ID AAS61273 standard; DNA; 5956 BP.
 XX
 AC AAS61273;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human gene regulation-associated gene oligonucleotide #228.
 XX
 KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preeclampsia; cardiac allograft vascular disease;

KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX Homo sapiens.
 OS
 PN WO200177375-A2.
 XX
 XX 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-EP003968.
 XX
 XX 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX
 XX
 DR WPI; 2002-017470/02.
 XX
 XX New nucleic acid sequences from chemically modified genes associated with
 PT gene regulation, useful for analyzing cytosine methylations for diagnosis
 PT and therapy of diseases e.g. severe combined immunodeficiency disease.
 XX
 XX Disclosure; SEQ ID NO 234; 26pp; English.
 XX
 XX The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The chemical
 CC pretreatment converts cytosine bases unmethylated at the 5-position to
 CC uracil or another base with hybridisation behaviour dissimilar to
 CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
 CC oligomers (or sets/arrays) and method are useful in the diagnosis of
 CC diseases (or predisposition to diseases) associated with gene regulation
 CC and in therapy of such diseases, by enabling analysis of the cytosine
 CC methylation patterns of such genes. Kits are provided. They are
 CC especially useful in diagnosis and therapy of e.g. severe combined
 CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
 CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
 CC syndrome, renal disease, preclampsia, graft versus-host disease. The
 CC present sequence is a sequence included in the sequence data for this
 CC specification and is associated with the human gene regulation-associated
 CC genes. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5956 BP; 1652 A; 171 C; 1433 G; 2700 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 6; Length 5956;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TTTTATATATATAGATAT 307
 DB 5381 TTTTATATATATAGATAT 5401

RESULT 7
 ABL07848
 ID ABL07848 standard; cDNA; 16739 BP.
 XX
 AC ABL07848;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18026.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX

OS Drosophila melanogaster.
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR P-PSDB; ABB63745.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 PT
 XX Claim 1; SEQ ID NO 18026; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 16739 BP; 4986 A; 3249 C; 3289 G; 5215 T; 0 U; 0 Other;

Query Match 1.3%; Score 21; DB 4; Length 16739;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 TAAATTAATAATTCATCTAT 229
 DB 2143 TAAATTAATAATTCATCTAT 2163

RESULT 8
 ABL07850
 ID ABL07850 standard; cDNA; 21828 BP.
 XX
 AC ABL07850;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18032.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PF
 XX 23-MAR-2001; 2001WO-US009231.
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI

XX WPI; 2001-656860/75.
 DR P-PSDB; ABB63747.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 18032; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 21828 BP; 5614 A; 4180 C; 4296 G; 6738 T; 0 U; 0 Other;
 Query Match 1.3%; Score 21; DB 4; Length 21828;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 TAAATTAAATTCATTCTAT 229
 Db |||||
 7232 TAAATTAAATTCATTCTAT 7252
 RESULT 9
 ABLN17338
 ID ABLN17338 standard; cDNA; 291 BP.
 AC ABLN17338;
 XX
 XX 24-JUN-2002 (first entry)
 XX
 DE Human ORFX polynucleotide sequence SEQ ID NO:3153.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW myasthenia gravis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200192523-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US010836.
 XX
 XX 30-MAY-2000; 2000US-0206132P.
 XX 29-AUG-2000; 2000US-0228716P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach MD;
 DR WPI; 2002-106308/14.
 DR P-PSDB; ABP01586.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX

PS Disclosure; SEQ ID NO 3153; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABLN15762 to ABLN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis.
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 291 BP; 107 A; 41 C; 56 G; 86 T; 0 U; 1 Other;
 Query Match 1.2%; Score 20; DB 6; Length 291;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 68 TCATTTCAGAGGATGAAAAA 87
 Db |||||
 164 TCATTTCAGAGGATGAAAAA 183
 RESULT 10
 AAL15883/c
 ID AAL15883 standard; cDNA; 785 BP.
 XX
 AC AAL15883;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 8340.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200151628-A2.
 XX
 XX 19-JUL-2001.
 XX
 XX 10-JAN-2001; 2001WO-US000798.
 XX
 XX 14-JAN-2000; 2000US-0176077P.
 XX 14-MAR-2000; 2000US-0189167P.
 XX 24-MAR-2000; 2000US-0192099P.
 XX 29-MAR-2000; 2000US-0193480P.
 XX 15-MAY-2000; 2000US-0205230P.
 XX 09-JUN-2000; 2000US-0211315P.
 XX 25-JUL-2000; 2000US-0220534P.
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Lillie J, Xu Y, Wang Y, Steinmann K;
 DR WPI; 2001-451856/48.
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT

```
XX PS Claim 1; Page 1509; 3695pp; English.
XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AA107544-AA126789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, monitoring, characterizing treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX CC activity
XX SQ Sequence 785 BP; 216 A; 172 C; 184 G; 200 T; 0 U; 13 Other;

Query Match 1.2%; Score 20; DB 4; Length 785;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTTTTCCTTAAGAAAAAAA 248
Db 709 TTTTTCCTTAAGAAAAAAA 690

RESULT 11
ADA49079
ID ADA49079 standard; DNA; 830 BP.
XX AC ADA49079;
XX DT 20-NOV-2003 (first entry)
XX DE Wheat gene conferring disease resistance in plants.
XX KW disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;
XX KW wheat.
XX OS Triticum aestivum.
XX PN WO2003000906-A2.
XX PD 03-JAN-2003.
XX PF 21-JUN-2002; 2002WO-IB002453.
XX PR 22-JUN-2001; 2001US-0300112P.
XX PR 26-SEP-2001; 2001US-0352277P.
XX PR 22-MAR-2002; 2002US-0366535P.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
XX PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX DR WPI; 2003-184052/18.
XX PT New polynucleotide comprising a plant nucleotide sequence having an open
XX PT reading frame that encodes a polypeptide associated with disease
XX PT resistance, useful for conferring resistance or tolerance to a plant
XX PT pathogen.
XX PS Disclosure; SEQ ID NO 1149; 299pp; English.
XX CC The invention relates to a novel isolated polynucleotide comprising a
XX CC plant nucleotide sequence having an open reading frame that encodes a
XX CC polypeptide associated with disease resistance or its fragment having
XX CC substantially the same activity as the full-length polypeptide. The
XX CC polynucleotide of the invention is useful for conferring resistance or
XX CC tolerance to a plant pathogen. The present sequence represents a gene
XX CC conferring disease resistance used in the invention.
XX SQ Sequence 830 BP; 177 A; 221 C; 219 G; 213 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 4; Length 785;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTTTTCCTTAAGAAAAAAA 248
Db 709 TTTTTCCTTAAGAAAAAAA 690

RESULT 11
ADA49079
ID ADA49079 standard; DNA; 830 BP.
XX AC ADA49079;
XX DT 20-NOV-2003 (first entry)
XX DE Wheat gene conferring disease resistance in plants.
XX KW disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;
XX KW wheat.
XX OS Triticum aestivum.
XX PN WO2003000906-A2.
XX PD 03-JAN-2003.
XX PF 21-JUN-2002; 2002WO-IB002453.
XX PR 22-JUN-2001; 2001US-0300112P.
XX PR 26-SEP-2001; 2001US-0352277P.
XX PR 22-MAR-2002; 2002US-0366535P.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
XX PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX DR WPI; 2003-184052/18.
XX PT New polynucleotide comprising a plant nucleotide sequence having an open
XX PT reading frame that encodes a polypeptide associated with disease
XX PT resistance, useful for conferring resistance or tolerance to a plant
XX PT pathogen.
XX PS Disclosure; SEQ ID NO 1149; 299pp; English.
XX CC The invention relates to a novel isolated polynucleotide comprising a
XX CC plant nucleotide sequence having an open reading frame that encodes a
XX CC polypeptide associated with disease resistance or its fragment having
XX CC substantially the same activity as the full-length polypeptide. The
XX CC polynucleotide of the invention is useful for conferring resistance or
XX CC tolerance to a plant pathogen. The present sequence represents a gene
XX CC conferring disease resistance used in the invention.
XX SQ Sequence 830 BP; 177 A; 221 C; 219 G; 213 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 8; Length 830;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1613 GATCGCGCCGAGCTCGCCA 1632
Db 545 GATCGCGCCGAGCTCGCCA 564

RESULT 12
ACC61661
ID ACC61661 standard; DNA; 1044 BP.
XX AC ACC61661;
XX DT 20-JUN-2003 (first entry)
XX DE Gene sequence #SEQ ID 2104.
XX KW Multiprotein complex; eukaryote; drug target; diagnosis; gene; ds.
XX OS Saccharomyces cerevisiae.
XX PN BP1258494-A1.
XX PD 20-NOV-2002.
XX PF 20-DEC-2001; 2001EP-00130253.
XX PR 15-MAY-2001; 2001EP-00111774.
XX PA (CELL-) CELLZOME AG.
XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;
XX PI Marzloch M, Schultz JD, Superti-Furga GD;
XX DR WPI; 2003-250078/25.
XX DR P-PSDB; ABR53619.
XX PT New isolated protein complexes useful for diagnosing a disease or
XX PT disorder, or as a target for an active agent of a pharmaceutical,
XX PT preferably a drug target in the treatment or prevention of disease or
XX PT disorder.
XX PS Disclosure; SEQ ID NO 2104; 17pp + Sequence Listing; English.
XX CC The invention relates to multiprotein complexes from eukaryotes. Proteins
XX CC of the invention and DNA sequences encoding them are given in records
XX CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are
XX CC obtainable by using a protein as a bait and isolating the set of proteins
XX CC which is attached thereto from cells. Such protein complexes may comprise
XX CC up to 30 distinct proteins. Protein complexes of the invention are useful
XX CC for diagnosing a disease or disorder, or as a target for an active agent
XX CC of a pharmaceutical, preferably a drug target in the treatment or
XX CC prevention of a disease or disorder. Note: The sequence data for this
XX CC patent is not represented in the printed specification, but is based on
XX CC sequence information supplied by the European Patent Office. The complete
XX CC document is available on CD-ROM
XX SQ Sequence 1044 BP; 367 A; 164 C; 229 G; 284 T; 0 U; 0 Other;

Query Match 1.2%; Score 20; DB 7; Length 1044;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TCATTTTCAGAGGATGAAAAA 87
Db 506 TCATTTTCAGAGGATGAAAAA 525

RESULT 13
AAZ52423/c
ID AAZ52423 standard; DNA; 1963 BP.
```

XX AAZ52423;
 AC 24-FEB-2000 (first entry)
 DT HTRM clone 2019742 DNA sequence.
 DE HTRM; human transcriptional regulatory molecule; arteriosclerosis; AIDS;
 XX arteriosclerosis; cirrhosis; cancer; leukaemia; diabetes mellitus; ss;
 KW Addison's disease; multiple sclerosis; rheumatoid arthritis; infection;
 KW trauma; myasthenia gravis; adenocarcinoma; immune disorder; treatment.
 XX Homo sapiens.
 OS W09957144-A2.
 PN 11-NOV-1999.
 XX 04-MAY-1999; 99WO-US009935.
 XX 05-MAY-1998; 98US-0084254P.
 PR 07-AUG-1998; 98US-0095827P.
 PR 02-OCT-1998; 98US-0102745P.
 XX (INCY-) INCYTE PHARM INC.
 PA Hillman JL, Bandman O, Lal P, Yue H, Reddy R, Tang YT;
 PI Gerstin EH, Patterson C, Baughn MR, Azimzai Y, Lu DAM;
 XX WPI; 2000-052941/04.
 DR P-PSDB; AAY73338.
 XX New peptides useful for diagnosis, prevention and treatment of cancer and
 PT immune disorders.
 XX Claim 9; Page 155; 193pp; English.
 PS AAZ52410-Z52474 are human transcriptional regulator molecule (HTRM)
 CC nucleotide sequences. The HTRM protein and nucleotide sequences are
 CC useful for preventing or treating disorders associated with decreased
 CC expression or activity of HTRM which include cell proliferative disorders
 CC such as arteriosclerosis and cirrhosis; cancers including adenocarcinoma
 CC and leukaemia; immune disorders such as AIDS, Addison's disease, diabetes
 CC mellitus, rheumatoid arthritis, multiple sclerosis, systemic lupus
 CC erythematosus, and myasthenia gravis; infections and trauma. Antagonists
 CC of the HTRM polypeptides are useful for treating or preventing disorders
 CC associated with increased expression or activity of HTRMs. HTRM
 CC screening libraries of immunogenic fragments or oligopeptides are useful for
 CC screening libraries of compounds in drug screening techniques.
 CC Polynucleotides encoding HTRM are useful for blocking the transcription
 CC of mRNA and regulating gene function by modulating the activity of HTRM.
 CC Vectors expressing HTRM or agonists can also be used to prevent or treat
 CC disorder associated with decreased HTRM expression. Antibodies which
 CC specifically bind HTRM and polynucleotides encoding HTRM are useful for
 CC diagnosing disorders associated with the expression of HTRM, particularly
 CC in assays that detect the expression of HTRM. Nucleotide sequences
 CC encoding HTRM may be useful to generate hybridization probes useful in
 CC mapping the naturally occurring genomic sequence and to detect
 CC differences in gene sequences among normal, carrier and affected
 CC individuals. Using diagnostic assays, cancer can be detected prior to the
 CC appearance of clinical symptoms and thereby progression of cancer can be
 CC prevented by aggressive treatment or preventive measures
 XX Sequence 1963 BP; 488 A; 499 C; 491 G; 485 T; 0 U; 0 Other;
 SQ Query Match 1.2%; Score 20; DB 3; Length 1963;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 TTAAGAAAAAACAGCCCA 255
 |||||
 1644 TTAAGAAAAAACAGCCCA 1625

RESULT 14
 AAI60732
 ID AAI60732 standard; cDNA; 1973 BP.
 XX AC AAI60732;
 XX 22-OCT-2001 (first entry)
 DT Human polynucleotide SEQ ID NO 4721.
 DE Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX Homo sapiens.
 OS MO200153312-A1.
 PN 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HVSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM41576.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PS Claim 1; SEQ ID NO 4721; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with neotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX Sequence 1973 BP; 485 A; 493 C; 505 G; 490 T; 0 U; 0 Other;
 SQ Query Match 1.2%; Score 20; DB 4; Length 1973;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 236 TTAAGAAAAAACAGCCCA 255

Db 315 TTAAGAAAAAACAGCCCA 334
 |||||
 RESULT 15
 AAI58946/C
 ID AAI58946 standard; cDNA; 1979 BP.
 XX
 AC AAI58946;
 CC
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1149.
 XX
 XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200153312-A1.
 PN
 XX 26-JUL-2001.
 DD
 XX 26-DEC-2000; 2000WO-US034263.
 PF
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM39790.
 DR
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PT
 XX Claim 1; SEQ ID NO 1149; 10078pp; English.
 XX
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AA442213) with neotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX Query 1979 BP; 498 A; 505 C; 485 G; 491 T; 0 U; 0 Other;
 SQ

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 13:02:10 ; Search time 144 Seconds
(without alignments)
6389.643 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcaggctcaatac.....ctcgttagtttggacggcg 1658

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	1.2	1979	4	US-09-620-312D-838
C 2	19	1.1	786431	4	US-09-751-389-3
C 3	19	1.1	1230025	4	US-09-198-452A-1
C 4	18	1.1	533	4	US-09-659-751-93
C 5	18	1.1	624	4	US-09-489-039A-4810
C 6	18	1.1	936	4	US-09-134-001C-1009
C 7	18	1.1	1428	4	US-09-328-352-1253
C 8	18	1.1	2311	2	US-08-712-709-6
C 9	18	1.1	2311	3	US-09-111-444-6
C 10	18	1.1	2311	3	US-09-541-228-6
C 11	18	1.1	2311	4	US-09-016-434-772
C 12	18	1.1	2370	4	US-09-031-295-1
C 13	18	1.1	3845	2	US-08-220-240A-4
C 14	18	1.1	3982	4	US-08-956-171E-520
C 15	18	1.1	4061	4	US-09-620-312D-363
C 16	18	1.1	5153	4	US-10-238-483-3
C 17	18	1.1	5658	4	US-09-647-344A-17
C 18	18	1.1	87350	3	US-08-781-891-79
C 19	18	1.1	87350	4	US-09-618-166-79
C 20	18	1.1	87543	4	US-09-791-211-3
C 21	18	1.1	4403765	3	US-09-103-840A-2
C 22	18	1.1	4411529	3	US-09-103-840A-1
C 23	17	1.0	243	1	US-08-341-568-2
C 24	17	1.0	243	2	US-08-911-020-2
C 25	17	1.0	285	4	US-09-134-001C-1300
C 26	17	1.0	291	4	US-09-184-418C-14
C 27	17	1.0	417	4	US-09-621-976-3318

C 28	17	1.0	498	4	US-09-252-991A-11412	Sequence 11412, A
C 29	17	1.0	621	4	US-09-107-532A-2883	Sequence 2883, Ap
C 30	17	1.0	666	1	US-08-463-115-4	Sequence 4, Appli
C 31	17	1.0	666	1	US-08-465-388-4	Sequence 4, Appli
C 32	17	1.0	714	4	US-09-540-236-757	Sequence 757, App
C 33	17	1.0	780	4	US-09-252-991A-3741	Sequence 3741, Ap
C 34	17	1.0	960	4	US-09-252-991A-3730	Sequence 3730, Ap
C 35	17	1.0	1071	4	US-09-252-991A-3633	Sequence 3633, Ap
C 36	17	1.0	1089	4	US-09-252-991A-3785	Sequence 3785, Ap
C 37	17	1.0	1209	3	US-09-105-537-21	Sequence 21, Appli
C 38	17	1.0	1278	4	US-09-252-991A-4932	Sequence 4932, Ap
C 39	17	1.0	1329	3	US-08-360-758-1	Sequence 1, Appli
C 40	17	1.0	1389	1	US-08-458-023B-1	Sequence 1, Appli
C 41	17	1.0	1389	3	US-09-111-556A-1	Sequence 1, Appli
C 42	17	1.0	1407	4	US-09-252-991A-11395	Sequence 11395, A
C 43	17	1.0	1594	2	US-08-955-713-1	Sequence 1, Appli
C 44	17	1.0	1717	4	US-09-634-137-3	Sequence 3, Appli
C 45	17	1.0	1857	3	US-09-299-378-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-620-312D-838/C
; Sequence 838, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: Pt_FL_Genes Version 1.0
; SEQ ID NO 838
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1053)
US-09-620-312D-838

Query Match 1.2%; Score 20; DB 4; Length 1979;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 TTAAGAAAAAACAAGCCCA 255
DB 1661 TTAAGAAAAAACAAGCCCA 1642

```
RESULT 2
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match 1.1%; Score 19; DB 4; Length 786431;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AAAGCAATATTTTAAATG 117
Db 620036 AAAGCAATATTTTAAATG 620018

RESULT 3
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6949
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
```

```
; NAME/KEY: misc_feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (150001)..(165000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (165001)..(180000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (180001)..(195000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (195001)..(210000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (210001)..(225000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (225001)..(240000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (240001)..(255000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (255001)..(270000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (270001)..(285000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (285001)..(300000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (300001)..(315000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (315001)..(330000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (330001)..(345000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (345001)..(360000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (360001)..(375000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (375001)..(390000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (390001)..(405000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (405001)..(420000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (420001)..(435000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (510001)..(525000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (525001)..(540000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (540001)..(555000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (555001)..(570000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (570001)..(585000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (585001)..(600000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (600001)..(615000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (615001)..(630000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (630001)..(645000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (645001)..(660000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (660001)..(675000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (675001)..(690000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (825001)..(840000)
```

```
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
```

Query Match 1.1%; Score 19; DB 4; Length 1230025;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1039 GAGTGTAGAGAGAACTGA 1057
      |||||
Db 1200765 GAGTGTAGAGAGAACTGA 1200783
```

```
RESULT 4
US-09-669-751-93
; Sequence 93, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 533
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-93
```

Query Match 1.1%; Score 18; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 689 CTTTGTGTTTGTGTTTCT 706
      |||||
Db 263 CTTTGTGTTTGTGTTTCT 280
```

```
RESULT 5
US-09-489-039A-4810
; Sequence 4810, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4810
```

```

; LENGTH: 624
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4810

Query Match
Best Local Similarity 1.1%; Score 18; DB 4; Length 624;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1378 CGCGTCGGCGCCGACGGG 1395
DB 394 CGCGTCGGCGCCGACGGG 411

RESULT 6
US-09-134-001C-1009
; Sequence 1009, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1009
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1009

Query Match
Best Local Similarity 1.1%; Score 18; DB 4; Length 936;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 CAAATATTTTAAATGAAT 120
DB 248 CAAATATTTTAAATGAAT 265

RESULT 7
US-09-328-352-1253
; Sequence 1253, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1253
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1253

Query Match
Best Local Similarity 1.1%; Score 18; DB 4; Length 1428;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 AAAAAACAGCCCATTAAG 260
DB 36 AAAAAACAGCCCATTAAG 53

RESULT 8

```

```

US-08-712-709-6/c
; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0118 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6

Query Match
Best Local Similarity 1.1%; Score 18; DB 2; Length 2311;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1501 CGGCTATAAAAGGCGGC 1518
DB 908 CGGCTATAAAAGGCGGC 891

RESULT 9
US-09-111-444-6/c
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```

; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-09-111-444-6

Query Match 1.1%; Score 18; DB 3; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

RESULT 10
US-09-541-228-6/c
; Sequence 6, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/541,228
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-09-111-444-6

Query Match 1.1%; Score 18; DB 3; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

RESULT 11
US-09-016-434-772/c
; Sequence 772, Application US/09016434
; Patent No. 650938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
; US-09-016-434-772

Query Match 1.1%; Score 18; DB 4; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

```

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-09-541-228-6

Query Match 1.1%; Score 18; DB 3; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

RESULT 11
US-09-016-434-772/c
; Sequence 772, Application US/09016434
; Patent No. 650938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
; US-09-016-434-772

Query Match 1.1%; Score 18; DB 4; Length 2311;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1501 CGGCTATAAAAAAGCGGC 1518
Db 908 CGGCTATAAAAAAGCGGC 891

```

RESULT 12
US-09-031-295-1/c
; Sequence 1, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGGER, Tubingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..1335
US-09-031-295-1
Query Match 1.1%; Score 18; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1501 CGGCTATAAAAGCGCGC 1518
Db 941 CGGCTATAAAAGCGCGC 924
RESULT 13
US-08-220-240A-4/c
; Sequence 4, Application US/08220240A
; Patent No. 5955291
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Matikainen, Marja-Terttu
; APPLICANT: Partanen, Juha
; APPLICANT: Makela, Tomi
; APPLICANT: Korhonen, Jaana
; TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR
; TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,240A
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FI93/00006
; FILING DATE: 08-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/817,800
; FILING DATE: 09-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/167,453
; FILING DATE: 15-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 29151/31958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..3450
US-08-220-240A-4
Query Match 1.1%; Score 18; DB 2; Length 3845;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1154 CACATAGGCTTCCTGGC 1171
Db 3390 CACATAGGCTTCCTGGC 3373
RESULT 14
US-08-956-171E-520
; Sequence 520, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

```
;
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/956,171E
;   FILING DATE: 20-Oct-1997
;   CLASSIFICATION: <unknown>
;
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 60/009,861
;   FILING DATE: January 5, 1996
;   APPLICATION NUMBER: 08/781,986
;   FILING DATE: January 3, 1997
;
; ATTORNEY/AGENT INFORMATION:
;   NAME: Mark J. Hyman
;   REGISTRATION NUMBER: 46,789
;   REFERENCE/DOCKET NUMBER: PB248P1
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (240) 314-1224
;     TELEFAX: (301) 309-8439
;
; INFORMATION FOR SEQ ID NO: 520:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 3982 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-08-956-171E-520

Query Match      1.1%; Score 18; DB 4; Length 3982;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43 TTTATAGTTAGATTACTT 60
DB      284 TTTATAGTTAGATTACTT 301

RESULT 15
US-09-620-312D-363
; Sequence 363, Application US/09620312D
; Patent NO. 659662
; GENERAL INFORMATION:
;   APPLICANT: Tang, Y. Tom
;   APPLICANT: Liu, Chenghua
;   APPLICANT: Asundi, Vinod
;   APPLICANT: Zhang, Jie
;   APPLICANT: Ren, Feiyan
;   APPLICANT: Chen, Rui-hong
;   APPLICANT: Zhao, Qing A.
;   APPLICANT: Wehrman, Tom
;   APPLICANT: Xue, Aidong J.
;   APPLICANT: Yang, Yonghong
;   APPLICANT: Wang, Jian-Rui
;   APPLICANT: Zhou, Ping
;   APPLICANT: Ma, Yungqing
;   APPLICANT: Wang, Dunrui
;   APPLICANT: Wang, Zhiwei
;   APPLICANT: John Tillinghast
;   APPLICANT: Drmanac, Radoje T.
;   TITLE OF INVENTION: No. 659662el Nucleic Acids and
;   TITLE OF INVENTION: Polypeptides
;   FILE REFERENCE: 784CIP2B
;   CURRENT APPLICATION NUMBER: US/09/620,312D
;   CURRENT FILING DATE: 2000-07-19
;   PRIOR APPLICATION NUMBER: 09/552,317
;   PRIOR FILING DATE: 2000-04-25
;   PRIOR APPLICATION NUMBER: 09/488,725
;   PRIOR FILING DATE: 2000-01-21
;   NUMBER OF SEQ ID NOS: 1105
;   SOFTWARE: pt_FL_genes Version 1.0
;   SEQ ID NO 363
;   LENGTH: 4061
;   TYPE: DNA
```

```
;
; ORGANISM: Homo sapiens
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (72)..(3062)
US-09-620-312D-363

Query Match      1.1%; Score 18; DB 4; Length 4061;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      38 TAAATTTTATAGTTAGAT 55
DB      3544 TAAATTTTATAGTTAGAT 3561

Search completed: September 2, 2004, 21:26:12
Job time : 152 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:41:35 ; Search time 783 Seconds
(without alignments)
10525.313 Million cell updates/sec

Title: US-10-732-721-1

Perfect score: 1658

Sequence: 1 gaattcaggctcaataac.....ctcgttagttgggacggcg 1658

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3267054 seqs, 2485319735 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*

14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	17	US-10-732-721-1
2	67	4.0	836	13	Sequence 1, Appli
3	66	4.0	918	13	Sequence 7252, Ap
4	40	2.4	894	13	Sequence 13779, A
5	22	1.3	2509	17	Sequence 3295, Ap
6	21	1.3	21	17	Sequence 17878, A
7	21	1.3	21	17	Sequence 5, Appli
8	21	1.3	5956	15	Sequence 5, Appli
9	20	1.2	1089	13	Sequence 234, App
10	20	1.2	1291	13	Sequence 6995, Ap
11	20	1.2	1291	16	Sequence 151164,
12	20	1.2	1690	15	Sequence 151164,
13	20	1.2	1924	13	Sequence 13, Appli
14	20	1.2	1924	16	Sequence 204099,

c 15	20	1.2	1979	15	US-10-037-270-838	Sequence 838, App
c 16	20	1.2	1979	16	US-10-117-722-838	Sequence 838, App
c 17	20	1.2	3255	16	US-10-094-749-1102	Sequence 1102, Ap
c 18	20	1.2	5823	15	US-10-240-453-256	Sequence 256, App
c 19	20	1.2	16163	10	US-09-764-891-6505	Sequence 6505, Ap
c 20	20	1.2	16163	13	US-10-091-414-270	Sequence 270, App
c 21	20	1.2	65454	17	US-10-293-864-11	Sequence 11, Appl
c 22	20	1.2	100301	17	US-10-450-826-83	Sequence 83, Appl
c 23	20	1.2	3186778	13	US-10-027-632-174961	Sequence 174961,
c 24	20	1.2	3186778	16	US-10-027-632-174961	Sequence 28761, A
c 25	19	1.1	399	16	US-10-369-493-29761	Sequence 34809,
c 26	19	1.1	485	13	US-10-027-632-324809	Sequence 324809,
c 27	19	1.1	485	16	US-10-027-632-324809	Sequence 312341,
c 28	19	1.1	559	13	US-10-027-632-312341	Sequence 312341,
c 29	19	1.1	559	16	US-10-027-632-312341	Sequence 71076, A
c 30	19	1.1	571	13	US-10-027-632-71076	Sequence 71076, A
c 31	19	1.1	571	16	US-10-027-632-71076	Sequence 14, Appl
c 32	19	1.1	604	13	US-10-001-885-14	Sequence 288247,
c 33	19	1.1	613	13	US-10-027-632-288247	Sequence 288248,
c 34	19	1.1	613	13	US-10-027-632-288248	Sequence 288247,
c 35	19	1.1	613	16	US-10-027-632-288247	Sequence 288248,
c 36	19	1.1	613	16	US-10-027-632-288248	Sequence 266112,
c 37	19	1.1	619	13	US-10-027-632-266112	Sequence 266113,
c 38	19	1.1	619	13	US-10-027-632-266112	Sequence 266112,
c 39	19	1.1	619	16	US-10-027-632-266112	Sequence 266113,
c 40	19	1.1	619	16	US-10-027-632-266112	Sequence 266113,
c 41	19	1.1	640	13	US-10-027-632-64828	Sequence 310725, A
c 42	19	1.1	640	13	US-10-027-632-310725	Sequence 64828, A
c 43	19	1.1	640	16	US-10-027-632-64828	Sequence 310725,
c 44	19	1.1	640	16	US-10-027-632-310725	Sequence 103276,
c 45	19	1.1	732	13	US-10-027-632-103276	

ALIGNMENTS

RESULT 1

US-10-732-721-1
; Sequence 1, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

Query Match	100.0%	Score 1658;	DB 17;	Length 1658;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1658;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTTATAGTTAGTACTT	60	
Db	1	GAATTCACGGCTCACAAATACAGTCATCTACATGTGATAAATTTTATAGTTAGTACTT	60	
Qy	61	CTTGTAATCATTTTACAGAGGATGAAAAAATCGCAAGAACCAATATTTTAATCAAT	120	
Db	61	CTTGTAATCATTTTACAGAGGATGAAAAAATCGCAAGAACCAATATTTTAATCAAT	120	
Qy	121	GATGCAATATACAAATTTAATTTACAAATTTATGTAAGATTACATTTGTTAGTTTCATAGA	180	
Db	121	GATGCAATATACAAATTTAATTTACAAATTTATGTAAGATTACATTTGTTAGTTTCATAGA	180	


```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13779
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-042-D8_FLI
US-10-425-114-13779

Query Match          4.0%; Score 66; DB 13; Length 918;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1554 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTTAGCAACG 1613
Db 1 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGTGTACGTTAGCAACG 60

QY 1614 ATCCGC 1619
Db 61 ATCCGC 66

RESULT 4
US-10-425-114-3295
; Sequence 3295, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3295
; LENGTH: 894
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700258323_FLI
US-10-425-114-3295

Query Match          2.4%; Score 40; DB 13; Length 894;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1580 GAACAGTAGTAGTCCCTGTGTACGTTAGCAACGATCCGC 1619
Db 1 GAACAGTAGTAGTCCCTGTGTACGTTAGCAACGATCCGC 40

RESULT 5
US-10-437-963-17878/c
; Sequence 17878, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 17878
; LENGTH: 2509
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23489C.1
US-10-437-963-17878

Query Match          1.3%; Score 22; DB 17; Length 2509;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 657 CGACGACGTGACGCTGATGCT 678
Db 568 CGACGACGTGACGCTGATGCT 547

RESULT 6
US-10-732-721-4
; Sequence 4, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: PAT_MRT4530_23489C.1
US-10-732-721-4

Query Match          1.3%; Score 21; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 ACTCCTTCCTCCGCTCCAG 1246
Db 1 ACTCCTTCCTCCGCTCCAG 21

RESULT 7
US-10-732-721-5/c
; Sequence 5, Application US/10732721
; Publication No. US20040163144A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-5

Query Match
Best Local Similarity 1.3%; Score 21; DB 17; Length 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1638 GCTCGTTAGTTGGGACGCG 1658
      |||||
Db 21 GCTCGTTAGTTGGGACGCG 1

RESULT 8
US-10-221-613-234
; Sequence 234, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 234
; LENGTH: 5956
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-234

Query Match
Best Local Similarity 1.3%; Score 21; DB 13; Length 5956;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 287 TTTTATATATAGATAT 307
      |||||
Db 5381 TTTTATATATAGATAT 5401

RESULT 9
US-10-198-846-6995/c
; Sequence 6995, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
```

```
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6995
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5, 6, 7, 10, 11, 12, 13, 16, 17, 18, 21, 22, 23, 28, 29,
; LOCATION: 30, 31, 33, 34, 35, 39, 41, 47, 49, 50, 53, 54, 56, 58,
; LOCATION: 59, 60, 61, 62, 63, 65, 66, 67, 69, 73, 74, 75, 79, 80,
; LOCATION: 82, 83, 84, 85, 86, 87, 88, 89, 92, 93, 94, 97, 99, 103,
; LOCATION: 107
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 109, 110, 111, 112, 113, 114, 115, 116, 124, 182, 202, 203,
; LOCATION: 204, 228, 406, 831, 870, 914, 921, 927, 929, 932, 947, 954,
; LOCATION: 972, 977, 980, 983, 994, 995, 997, 1017, 1032, 1043, 1050,
; LOCATION: 1053, 1081, 1083, 1085, 1088
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-6995

Query Match
Best Local Similarity 1.2%; Score 20; DB 15; Length 1089;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 TTTTCTTATAGAAAAAAA 248
      |||||
Db 867 TTTTCTTATAGAAAAAAA 848

RESULT 10
US-10-027-632-151164
; Sequence 151164, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151164
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151164

Query Match
Best Local Similarity 1.2%; Score 20; DB 13; Length 1291;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1578 AAGAACAGTAGTAGCCCTG 1597
      |||||
Db 276 AAGAACAGTAGTAGCCCTG 295
```

RESULT 11
US-10-027-632-151164
; Sequence 151164, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151164
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-151164

Query Match 1.2%; Score 20; DB 16; Length 1291;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1578 AAGAACAGTAGTAGTCCTG 1597
|||||
Db 276 AAGAACAGTAGTAGTCCTG 295

RESULT 12
US-10-268-822-13/c
; Sequence 13, Application US/10268822
; Publication No. US20030150004A1
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; TITLE OF INVENTION: Metabolism
; FILE REFERENCE: P02729US2
; CURRENT APPLICATION NUMBER: US/10/268,822
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Mouse
US-10-268-822-13

Query Match 1.2%; Score 20; DB 15; Length 1690;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 TGAGTTCAACATCTTTCTTA 919
|||||
Db 1227 TGAGTTCAACATCTTTCTTA 1208

RESULT 13
US-10-027-632-204099/c
; Sequence 204099, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204099
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1924)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-204099

Query Match 1.2%; Score 20; DB 13; Length 1924;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TTTCCTTAAGAAAAAACA 250
|||||
Db 1770 TTTCCTTAAGAAAAAACA 1751

RESULT 14
US-10-027-632-204099/c
; Sequence 204099, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204099
; LENGTH: 1924
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1924)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-204099

Query Match 1.2%; Score 20; DB 16; Length 1924;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TTTCCTTAAGAAAAAACA 250
|||
Db 1770 TTTCCTTAAGAAAAAACA 1751

RESULT 15

US-10-037-270-838/c
; Sequence 838, Application US/10037270
; Publication NO. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 838
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1053)
US-10-037-270-838

Query Match 1.2%; Score 20; DB 15; Length 1979;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 TTAAGAAAAAACAGCCCA 255
|||
Db 1661 TTAAGAAAAAACAGCCCA 1642

Search completed: September 2, 2004, 21:49:23
Job time : 788 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 12:48:50 ; Search time 4953 Seconds
(without alignments)
11778.434 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcacggctcacataac.....ctcgtagttgggacggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 37577330 seqs, 17593059518 residues

Word size : 0

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents NA Main:*

1: /cgn2_6/ptodata/2/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US080 COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092 COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093 COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095 COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096 COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098 COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Best Local Similarity 100.0%; Pred. No. 0;				Matches 1658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Result No.	Score	Query Match	Length	DB ID	Description		
1	1658	100.0	1658	100	US-60-434-242-1	Sequence 1, Appl	
2	1658	100.0	1658	110	US-60-531-039-16	Sequence 16, Appl	
3	79	4.8	385	26	US-09-620-111B-7324	Sequence 7324, Ap	
4	68	4.1	308	18	US-09-304-517A-40406	Sequence 40406, A	
5	68	4.1	308	19	US-09-371-146A-40406	Sequence 40406, A	
6	68	4.1	308	42	US-09-985-678-40406	Sequence 40406, A	
7	67	4.0	396	35	US-09-865-439A-30469	Sequence 30469, A	
8	67	4.0	396	76	US-60-207-458-74766	Sequence 74766, A	
9	67	4.0	470	35	US-09-865-439A-27134	Sequence 27134, A	
10	67	4.0	470	76	US-60-207-458-71431	Sequence 71431, A	
11	67	4.0	836	46	US-10-155-881-37309	Sequence 37309, A	
12	67	4.0	836	51	US-10-425-114-7252	Sequence 7252, Ap	
13	67	4.0	836	51	US-10-425-114A-7252	Sequence 7252, Ap	
14	67	4.0	836	88	US-60-312-544-973	Sequence 973, App	
15	66	4.0	861	27	US-09-654-617-273716	Sequence 273716,	
16	66	4.0	861	29	US-09-684-016-273716	Sequence 13779, A	
17	66	4.0	918	51	US-10-425-114-13779	Sequence 13779, A	
18	66	4.0	918	51	US-10-425-114A-13779	Sequence 13779, A	
19	64	3.9	204	18	US-09-304-517A-39577	Sequence 39577, A	
20	64	3.9	204	19	US-09-371-146A-39577	Sequence 39577, A	
21	64	3.9	204	42	US-09-985-678-39577	Sequence 39577, A	
22	60	3.6	527	35	US-09-865-439A-48132	Sequence 48132, A	
23	60	3.6	527	76	US-60-207-458-92395	Sequence 92395, A	
24	59	3.6	280	18	US-09-303-031A-3994	Sequence 3994, Ap	
25	59	3.6	280	18	US-09-304-517A-57027	Sequence 57027, A	
26	59	3.6	280	19	US-09-371-146A-57027	Sequence 57027, A	
27	59	3.6	280	36	US-09-894-949-3994	Sequence 3994, Ap	
28	59	3.6	280	36	US-09-894-949A-3994	Sequence 3994, Ap	
29	59	3.6	280	42	US-09-985-678-57027	Sequence 57027, A	
30	52	3.1	437	18	US-09-304-517A-79013	Sequence 79013, A	
31	52	3.1	437	19	US-09-371-146A-79013	Sequence 79013, A	
32	52	3.1	437	24	US-09-565-306-1774	Sequence 1774, Ap	
33	52	3.1	437	42	US-09-985-678-79013	Sequence 79013, A	
34	50	3.0	460	35	US-09-865-439A-43255	Sequence 43255, A	
35	50	3.0	460	76	US-60-207-458-87528	Sequence 87528, A	
36	47	2.8	686	35	US-09-865-439A-31938	Sequence 31938, A	
37	47	2.8	686	76	US-60-207-458-76235	Sequence 76235, A	
38	43	2.6	315	18	US-09-304-517A-78600	Sequence 78600, A	
39	43	2.6	315	19	US-09-371-146A-78600	Sequence 78600, A	
40	43	2.6	315	24	US-09-565-306-3420	Sequence 3420, Ap	
41	43	2.6	315	42	US-09-985-678-78600	Sequence 78600, A	
42	42	2.5	316	35	US-09-865-439A-44537	Sequence 44537, A	
43	42	2.5	316	76	US-60-207-458-88800	Sequence 88800, A	
44	40	2.4	291	18	US-09-304-517A-39265	Sequence 39265, A	
45	40	2.4	291	19	US-09-371-146A-39265	Sequence 39265, A	

ALIGNMENTS

RESULT 1
US-60-434-242-1
; Sequence 1, Application US/60434242
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/60/434, 242
; CURRENT FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-60-434-242-1
Query Match 100.0%; Score 1658; DB 100; Length 1658;


```
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 40406
; LENGTH: 308
; TYPE: DNA
; ORGANISM: Zea mays
US-09-985-678-40406

Query Match
  4.1%; Score 68; DB 42; Length 308;
Best Local Similarity 100.0%; Pred. No. 4e-23;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1552 AGCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 1611
    |||||||
Db 1 AGCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 60

QY 1612 CGATCCGC 1619
    |||||||
Db 61 CGATCCGC 68

RESULT 7
US-09-865-439A-30469
; Sequence 30469, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 30469
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-052-P1-K1-G5
US-09-865-439A-30469

Query Match
  4.0%; Score 67; DB 35; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 1612
    |||||||
Db 1 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 60

QY 1613 GATCCGC 1619
    |||||||
Db 61 GATCCGC 67

RESULT 8
US-60-207-458-74766
; Sequence 74766, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
```

```
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 74766
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-052-P1-K1-G5
US-60-207-458-74766

Query Match
  4.0%; Score 67; DB 76; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 1612
    |||||||
Db 1 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 60

QY 1613 GATCCGC 1619
    |||||||
Db 61 GATCCGC 67

RESULT 9
US-09-865-439A-27134
; Sequence 27134, Application US/09865439A
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51936)B
; CURRENT APPLICATION NUMBER: US/09/865,439A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,458
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 119126
; SEQ ID NO 27134
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(470)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3354-008-P1-K1-B11
US-09-865-439A-27134

Query Match
  4.0%; Score 67; DB 35; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 1612
    |||||||
Db 1 GCAGCAACAGCCCAACACCTAAAGAACAGTAGTCCCTGTGTAGCTTAGCAA 60

QY 1613 GATCCGC 1619
    |||||||
Db 61 GATCCGC 67

RESULT 10
US-60-207-458-71431
; Sequence 71431, Application US/60207458
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Deikman, Jill
; APPLICANT: Hardeman, Kristine J.
; APPLICANT: La Rosa, Thomas J.
```

; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ruan, Yijun G.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Sammons, R. Douglas
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Wu, Kunsheng
; APPLICANT: Xu, Nanfei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(51936)A
; CURRENT APPLICATION NUMBER: US/60/207,458
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 152403
; SEQ ID NO 71431
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB3354-008-P1-K1-B11
US-10-732-721-1.oligo.rnpsm

Query Match 4.0%; Score 67; DB 76; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 11
US-10-155-881-37309
; Sequence 37309, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyeva, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 37309
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
US-10-155-881-37309

Query Match 4.0%; Score 67; DB 46; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 12
US-10-425-114-7252
; Sequence 7252, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114-7252

Query Match 4.0%; Score 67; DB 51; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 13
US-10-425-114A-7252
; Sequence 7252, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7252
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-10-425-114A-7252

Query Match 4.0%; Score 67; DB 51; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1553 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 1612
Db 1 GCAGCAACAGCCAAACACCTAAAGAACAGTAGTAGTCCTCGTGTACGCTCTAGCAAC 60
QY 1613 GATCCGC 1619
Db 61 GATCCGC 67

RESULT 14
US-60-312-544-973
; Sequence 973, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

```

; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 973
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (108)..(449)
; OTHER INFORMATION: Clone ID: 700623908_FLI
US-60-312-544-973

```

```

Query Match          4.0%; Score 67; DB 88; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 GCAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGTACGCTTAGCAAC 1612
Db      1 GCAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGTACGCTTAGCAAC 60

QY 1613 GATCCGC 1619
Db      61 GATCCGC 67

```

```

RESULT 15
US-09-654-617-273716
; Sequence 273716, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 273716
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-273716

```

```

Query Match          4.0%; Score 66; DB 27; Length 861;
Best Local Similarity 100.0%; Pred. No. 4.4e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1554 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGTACGCTTAGCAACG 1613
Db      24 CAGCAACAGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGTACGCTTAGCAACG 83

QY 1614 ATCCGC 1619
Db      84 ATCCGC 89

```

Search completed: September 2, 2004, 21:23:30
Job time : 4955 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 13:36:00 ; Search time 586 Seconds
(without alignments)
9801.724 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcaggctcacataac.....ctcgttagttgggacggcg 1658

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 5837357 seqs, 1732150321 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11674714

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents NA New:

- 1: /cgn2_6/ptodata/1/pna/PCT NEW COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06 NEW COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07 NEW COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08 NEW COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09 NEW COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US60 NEW COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1658	100.0	1658	7	US-10-732-721-1
2	69	4.2	975	6	US-10-425-115-181022
3	40	2.4	1034	6	US-10-425-115-181033
C 4	38	2.3	528	6	US-10-425-115-47987
C 5	37	2.2	2211	6	US-10-425-115-82158
C 6	23	1.4	868	6	US-10-425-115-68623
C 7	21	1.3	21	7	US-10-732-721-4
C 8	21	1.3	21	7	US-10-732-721-5
C 9	21	1.3	772	6	US-10-425-115-30821
C 10	21	1.3	88344	6	US-10-918-754-16910
C 11	21	1.3	24289	6	US-10-918-754-16866
C 12	20	1.2	291	5	US-09-865-590A-3153
C 13	20	1.2	751	6	US-10-425-115-12790
C 14	20	1.2	955	6	US-10-425-115-163648
C 15	20	1.2	1690	1	PCT-US03-32273-13
C 16	20	1.2	1973	1	PCT-US04-07412-1190
C 17	20	1.2	1973	8	US-10-389-559-1190
C 18	20	1.2	12988	6	US-10-472-965-729
C 19	20	1.2	12988	6	US-10-896-164-8392
C 20	20	1.2	12988	6	US-10-896-164-11776
C 21	20	1.2	12988	8	US-10-868-184A-8392
C 22	20	1.2	12988	8	US-10-868-184A-11776
C 23	20	1.2	12988	8	US-10-868-184-8392
C 24	20	1.2	12988	8	US-10-868-184-11776

25	20	1.2	948061	6	US-10-915-727-12219	Sequence 12219, A
26	19	1.1	201	9	US-60-568-845-11203	Sequence 11203, A
27	19	1.1	201	9	US-60-568-845-11242	Sequence 11242, A
C 28	19	1.1	312	7	US-10-674-124A-17773	Sequence 17773, A
29	19	1.1	581	6	US-10-425-115-120930	Sequence 120930, A
30	19	1.1	601	5	US-09-949-003C-25389	Sequence 25389, A
31	19	1.1	601	5	US-09-949-003C-25390	Sequence 25390, A
32	19	1.1	601	5	US-09-949-003C-25391	Sequence 25391, A
33	19	1.1	601	5	US-09-949-003C-28894	Sequence 28894, A
34	19	1.1	601	5	US-09-949-003C-28895	Sequence 28895, A
35	19	1.1	601	5	US-09-949-003C-28896	Sequence 28896, A
C 36	19	1.1	754	7	US-10-767-701-8179	Sequence 8179, Ap
C 37	19	1.1	835	6	US-10-425-115-146524	Sequence 146524, A
C 38	19	1.1	949	9	US-60-579-062-25357	Sequence 25357, A
39	19	1.1	1329	6	US-10-425-115-94611	Sequence 94611, A
C 40	19	1.1	1484	1	PCT-US04-05654-2721	Sequence 2721, Ap
41	19	1.1	8614	6	US-10-896-164-8129	Sequence 8129, Ap
42	19	1.1	8614	8	US-10-868-184A-8129	Sequence 8129, Ap
43	19	1.1	8614	8	US-10-868-184-8129	Sequence 8129, Ap
44	19	1.1	14575	9	US-60-563-440-12172	Sequence 12172, A
C 45	19	1.1	17493	9	US-60-563-440-12170	Sequence 12170, A

ALIGNMENTS

RESULT 1

```

US-10-732-721-1
; Sequence 1, Application US/10732721
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
; FILE REFERENCE: 38-15(52826)A
; CURRENT APPLICATION NUMBER: US/10/732,721
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/60/434,242
; PRIOR FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1658
; TYPE: DNA
; ORGANISM: Zea mays
US-10-732-721-1

```

Query Match		100.0%	Score 1658;	DB 7;	Length 1658;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1658;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GAATTCAGGCTCA	CAATPACCA	GTCACTACAT	GTGATPAAATTTTATAGTTAGATTACTT 60
Db	1	GAATTCAGGCTCA	CAATPACCA	GTCACTACAT	GTGATPAAATTTTATAGTTAGATTACTT 60
Qy	61	CTTGTAACTATT	CAGAGGATGA	AAAAAATCC	GAAGAAGCAAAATTTTAAATGAAT 120
Db	61	CTTGTAACTATT	CAGAGGATGA	AAAAAATCC	GAAGAAGCAAAATTTTAAATGAAT 120
Qy	121	GATGCAATATACA	AAATTTAA	TATACACAA	TATATGAAGATTACATTTTCTTTTCTATAGA 180
Db	121	GATGCAATATACA	AAATTTAA	TATACACAA	TATATGAAGATTACATTTTCTTTTCTATAGA 180
Qy	181	AATCAATTTCTA	GTAGTCAATA	ATATGCTAA	ATTTAAATTTCTATTTTCTTTTCTTAAAG 240
Db	181	AATCAATTTCTA	GTAGTCAATA	ATATGCTAA	ATTTAAATTTCTATTTTCTTTTCTTAAAG 240
Qy	241	AAAAAAACAGCC	CAATTAAG	GACCATTA	GAATGGCGCTGCTCCATTTTCTTTTATATA 300
Db	241	AAAAAAACAGCC	CAATTAAG	GACCATTA	GAATGGCGCTGCTCCATTTTCTTTTATATA 300
Qy	301	GAGATATGAGTT	GTGTCTACT	AGCCACAT	CCCTCCATGCGTGTCTTTTATTACATG 360
Db	301	GAGATATGAGTT	GTGTCTACT	AGCCACAT	CCCTCCATGCGTGTCTTTTATTACATG 360

QY 361 AAAAAATGACGGCTGCTTTTGCATTAAGGGTCTAGAGATGACGATGCAACGCTTCGT 420
 Db 361 AAAAAATGACGGCTGCTTTTGCATTAAGGGTCTAGAGATGACGATGCAACGCTTCGT 420
 QY 421 ATTCACTGACAGTAAATTTTAAATGACGGAGGCGCTTTGATTCCCTCAGCAGCT 480
 Db 421 ATTCACTGACAGTAAATTTTAAATGACGGAGGCGCTTTGATTCCCTCAGCAGCT 480
 QY 481 CTGCGCGCCGCTGCTTCGTGTACGGCAGTGACGAGGACGAGCAGCTCTACGTGCCACG 540
 Db 481 CTGCGCGCCGCTGCTTCGTGTACGGCAGTGACGAGGACGAGCAGCTCTACGTGCCACG 540
 QY 541 TGCCTGGCTGCTGCGCCACCGGTGATGCAAGGCTTGTCTGCTGCGTGGCTGGC 600
 Db 541 TGCCTGGCTGCTGCGCCACCGGTGATGCAAGGCTTGTCTGCTGCGTGGCTGGC 600
 QY 601 AGCGAGAGGTTGAAGCTACGGCGTGGCTGCTGCGCGCTGCTAGCACGCGAC 660
 Db 601 AGCGAGAGGTTGAAGCTACGGCGTGGCTGCTGCGCGCTGCTAGCACGCGAC 660
 QY 661 GACGTGACGCTGATGCTAGCGTGTCTACTTTTGTGTTTCTGCTGCTGCGATG 720
 Db 661 GACGTGACGCTGATGCTAGCGTGTCTACTTTTGTGTTTCTGCTGCTGCGATG 720
 QY 721 ATGCAGTAAATTAACGCTGTATCTCATGTGATTCATGACGCTGTGCGGATTTCAA 780
 Db 721 ATGCAGTAAATTAACGCTGTATCTCATGTGATTCATGACGCTGTGCGGATTTCAA 780
 QY 781 ATCAGTAAATCACAATGTCATCCCATCGACTTGTGTCGGCGATAAATCTGCTAGG 840
 Db 781 ATCAGTAAATCACAATGTCATCCCATCGACTTGTGTCGGCGATAAATCTGCTAGG 840
 QY 841 ATTTGTCTTTTCAATGATGCTTGAATAGAGAGGAAATTAATATTAACATAATAT 900
 Db 841 ATTTGTCTTTTCAATGATGCTTGAATAGAGAGGAAATTAATATTAACATAATAT 900
 QY 901 GAGTTCAACATCTTCTTAATACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Db 901 GAGTTCAACATCTTCTTAATACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 QY 961 TAACGTGTTAAGAACCCCTGTTACCGAGTGTAGTCAAGGCTGCTTAACTGAATATAG 1020
 Db 961 TAACGTGTTAAGAACCCCTGTTACCGAGTGTAGTCAAGGCTGCTTAACTGAATATAG 1020
 QY 1021 TAGAACACAAAGGAGCAAGAGTGTAGAGGAACTGATTTCTTGTACTATATGTTGCTG 1080
 Db 1021 TAGAACACAAAGGAGCAAGAGTGTAGAGGAACTGATTTCTTGTACTATATGTTGCTG 1080
 QY 1081 CTCTCCAAAGGTTACATGATGAGGATCTCTCTATTTATAGACAAAGTGGTTT 1140
 Db 1081 CTCTCCAAAGGTTACATGATGAGGATCTCTCTATTTATAGACAAAGTGGTTT 1140
 QY 1141 CAGGCATATGGCCACATAGGCTTCTGCGCCCAAGAAAGGTTTCTTAACTACCATCT 1200
 Db 1141 CAGGCATATGGCCACATAGGCTTCTGCGCCCAAGAAAGGTTTCTTAACTACCATCT 1200
 QY 1201 TCAGGCGCGGTGTGGTGTCTTCACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 1201 TCAGGCGCGGTGTGGTGTCTTCACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 QY 1261 GGGTCTACGCTGACGCGCAGCATGTGCGCGAGGGGATCTCTGCGCGGGAATGCG 1320
 Db 1261 GGGTCTACGCTGACGCGCAGCATGTGCGCGAGGGGATCTCTGCGCGGGAATGCG 1320
 QY 1321 CGCAGGCGCATCGCTCGACAGCAGTACGTGTGCGCCCGCGAGGCTCTCTGCGACACG 1380
 Db 1321 CGCAGGCGCATCGCTCGACAGCAGTACGTGTGCGCCCGCGAGGCTCTCTGCGACACG 1380
 QY 1381 GTCCGCGCGCAGCGGCGGAGGACCTTGCATTCGTCGCGTGGCTGACACGCTCCCT 1440
 Db 1381 GTCCGCGCGCAGCGGCGGAGGACCTTGCATTCGTCGCGTGGCTGACACGCTCCCT 1440

QY 1441 GGTCCGGCTTCCGGCTTTGCTGCTGCGCGCGAGTGTCCGCTCTCCCCACCGGA 1500
 Db 1441 GGTCCGGCTTCCGGCTTTGCTGCTGCGCGCGAGTGTCCGCTCTCCCCACCGGA 1500
 QY 1501 CGGCTATAAAAGGCGCGCCACCTGATCTCCATCTCACAAAGCAAGCAGCAGCAAC 1560
 Db 1501 CGGCTATAAAAGGCGCGCCACCTGATCTCCATCTCACAAAGCAAGCAGCAGCAAC 1560
 QY 1561 AGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTAGCAACGATCCGG 1620
 Db 1561 AGCCACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTAGCAACGATCCGG 1620
 QY 1621 CGAGCTCCCGAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658
 Db 1621 CGAGCTCCCGAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1658

RESULT 2

US-10-425-115-181022
 ; Sequence 181022, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 181022
 ; LENGTH: 975
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_9666C.1
 US-10-425-115-181022

Query Match 4.2%; Score 69; DB 6; Length 975;

Best Local Similarity 100.0%; Pred. No. 3.9e-25;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1551 CAGCAGCAACAGCAACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTAGCA 1610
 Db 21 CAGCAGCAACAGCAACACCTAACTAAAGAACAGTAGTAGTCCCTGCTGCTAGCA 80
 QY 1611 ACGATCCGC 1619
 Db 81 ACGATCCGC 89

RESULT 3

US-10-425-115-181033
 ; Sequence 181033, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 181033
 ; LENGTH: 1034
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_9667C.1
 US-10-425-115-181033

Query Match 2.4%; Score 40; DB 6; Length 1034;
 Best Local Similarity 100.0%; Pred. No. 5.9e-10;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1580 GAACAGTAGTACGCTCGTGTACGCTAGCAACGATCCGC 1619
 Db 1 GAACAGTAGTACGCTCGTGTACGCTAGCAACGATCCGC 40

RESULT 4
 US-10-425-115-47987/c
 ; Sequence 47987, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 47987
 ; LENGTH: 528
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(528)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_14376C.1
 US-10-425-115-47987

Query Match 2.3%; Score 38; DB 6; Length 528;
 Best Local Similarity 100.0%; Pred. No. 6.4e-09;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 944 TAGAGCCAGTCGCTGATACCTGTGTAAGAACCCCTTG 981
 Db 109 TAGAGCCAGTCGCTGATACCTGTGTAAGAACCCCTTG 72

RESULT 5
 US-10-425-115-82158/c
 ; Sequence 82158, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 82158
 ; LENGTH: 2211
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_174942C.1
 US-10-425-115-82158

Query Match 2.2%; Score 37; DB 6; Length 2211;
 Best Local Similarity 100.0%; Pred. No. 2.3e-08;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1088 AAGTTTACATGATAGGGATCTCTCTATTATTATA 1124
 Db 1 AAGTTTACATGATAGGGATCTCTCTATTATTATA 1124

Db 2209 AAGTTTACATGATAGGGATCTCTCTATTATTATA 2173

RESULT 6
 US-10-425-115-68623/c
 ; Sequence 68623, Application US/10425115
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 68623
 ; LENGTH: 868
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_162578C.1
 US-10-425-115-68623

Query Match 1.4%; Score 23; DB 6; Length 868;
 Best Local Similarity 100.0%; Pred. No. 0.46;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 AGTCATCTACATGTGATAAATT 44
 Db 535 AGTCATCTACATGTGATAAATT 513

RESULT 7
 US-10-732-721-4
 ; Sequence 4, Application US/10732721
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology, LLC
 ; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
 ; FILE REFERENCE: 38-15(52826)A
 ; CURRENT APPLICATION NUMBER: US/10/732,721
 ; CURRENT FILING DATE: 2003-12-10
 ; PRIOR APPLICATION NUMBER: US/60/434,242
 ; PRIOR FILING DATE: 2002-12-18
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4
 ; LENGTH: 21
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-10-732-721-4

Query Match 1.3%; Score 21; DB 7; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1226 ACTCCTTCCCTCCGCTCCAG 1246
 Db 1 ACTCCTTCCCTCCGCTCCAG 21

RESULT 8
 US-10-732-721-5/c
 ; Sequence 5, Application US/10732721
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology, LLC
 ; TITLE OF INVENTION: Maize Embryo-Specific Promoter Compositions and Methods for Use
 ; FILE REFERENCE: 38-15(52826)A
 ; CURRENT APPLICATION NUMBER: US/10/732,721
 ; CURRENT FILING DATE: 2003-12-10

US-10-918-754-16910

RESULT 13

```
US-10-425-115-12790
; Sequence 12790, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 12790
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_111661C.1
US-10-425-115-12790

Query Match          1.2%; Score 20; DB 6; Length 751;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 958 TGATAACGCTGTTAAGAACCC 977
      ||||||||||||||||||
Db 723 TGATAACGCTGTTAAGAACCC 742

RESULT 14
US-10-425-115-163648
; Sequence 163648, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 163648
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80823C.1
US-10-425-115-163648

Query Match          1.2%; Score 20; DB 6; Length 955;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 958 TGATAACGCTGTTAAGAACCC 977
      ||||||||||||||||||
Db 883 TGATAACGCTGTTAAGAACCC 902

RESULT 15
PCT-US03-32273-13/c
; Sequence 13, Application PC/TUS0332273
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Wei, Ping
; APPLICANT: Chua, Steven
; TITLE OF INVENTION: Screening Systems and Methods for Identifying Modulators of Xenob
; FILE REFERENCE: P02729W02
; CURRENT APPLICATION NUMBER: PCT/US03/32273
; CURRENT FILING DATE: 2003-10-09
```

```
; PRIOR APPLICATION NUMBER: PCT/US 01/29672
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 10/268,822
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 10/219,590
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Mouse
PCT-US03-32273-13

Query Match          1.2%; Score 20; DB 1; Length 1690;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 TCAGTTCAACATCTTTCTTA 919
      ||||||||||||||||||
Db 1227 TCAGTTCAACATCTTTCTTA 1208

Search completed: September 2, 2004, 21:36:05
Job time : 588 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 10:46:24 ; Search time 3928 Seconds
(without alignments)
12604.760 Million cell updates/sec

Title: US-10-732-721-1
Perfect score: 1658
Sequence: 1 gaattcacggtcacatac.....ctcgttagtttggagcggcg 1658

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estmd:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_est1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:
29: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	451	27.2	779	28	CC418961
C 2	436	26.3	728	29	CG152000
C 3	374	22.6	498	29	CG151999
C 4	313	18.9	754	29	CG146202

5	305	18.4	938	29	CG146205
6	281	16.9	373	29	CC654144
7	248	15.0	949	28	BZ705097
C 8	86	5.2	728	29	CG164746
C 9	67	4.0	786	29	CG062076
C 10	67	4.0	858	29	CG062075
C 11	64	3.9	438	29	CG193723
C 12	64	3.9	446	29	CG193725
C 13	64	3.9	732	28	CC352433
C 14	64	3.9	884	28	CC006284
C 15	60	3.6	844	29	CG096594
C 16	60	3.6	914	29	CG124360
C 17	60	3.6	925	29	CG096596
C 18	59	3.6	870	11	AY106704
C 19	56	3.4	879	29	CG118361
C 20	55	3.3	740	29	CG041830
C 21	48	2.9	656	28	BH872036
C 22	48	2.9	775	28	BZ961663
C 23	48	2.9	826	28	BZ988278
C 24	46	2.8	862	29	CG175932
C 25	46	2.8	985	29	CG175933
C 26	44	2.7	866	29	CG118362
C 27	38	2.3	528	10	AW216120
C 28	38	2.3	794	28	BZ991151
C 29	37	2.2	574	10	AW155643
C 30	37	2.2	665	28	CC443358
C 31	34	2.1	754	28	BZ828297
C 32	34	2.1	776	29	CG113123
C 33	34	2.1	802	28	CC382778
C 34	34	2.1	807	28	BZ828291
C 35	34	2.1	949	29	CG041829
C 36	34	2.1	993	29	CC463184
C 37	30	1.8	225	28	BZ750783
C 38	30	1.8	955	29	CG097258
C 39	29	1.7	656	29	CG050059
C 40	29	1.7	656	29	CG050060
C 41	29	1.7	695	29	CC726211
C 42	28	1.7	856	29	CC266334
C 43	28	1.7	910	29	CC615699
C 44	28	1.7	949	29	CG266345
C 45	26	1.6	1032	29	CG046479

ALIGNMENTS

RESULT 1
LOCUS CC418961/c
DEFINITION PUEDX81TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra230M18,
genomic survey sequence.
ACCESSION CC418961
VERSION CC418961.1 GI:30899051
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 779)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw
TTGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

CC418961 779 bp DNA linear GSS 19-MAY-2003
PUEDX81TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra230M18,

CG146205 PUCV14TD
CC654144 OGUDK20TV
BZ705097 PUCB143TD
CG164746 PU1U51TB
CG062076 PUPD96TD
CG062075 PUPD96TB
CG193723 PUDJ37TB
CG193725 PUDJ37TD
CC352433 PUEHX59TD
CC006284 PUD1N46TD
CG096594 PUFUD53TB
CG124360 PUFUD53TB
CG096596 PUFUD53TD
AY106704 Zea mays
CG118361 PUFVU43TB
CG041830 PUFW117TD
BH872036 h38d06_b
BZ961663 PUBND60TD
BZ988278 PUBNE60TD
CG175932 PUFY251TB
CG175933 PUFY251TD
AW216120 687029C03
CG113123 PUGY13TD
CG382778 PUBIE27TD
BZ828291 PUGAG16TB
CG041829 PUFW117TB
CC463184 ZMMBGC034
BZ750783 PUFED34TD
CG097258 PUFWE84TB
CG050059 PUFUA36TB
CG050060 PUFUA36TD
CC726211 ZMMBB0013
CC266334 OGIAB48TH
CC615699 OGHP60TV
CG266345 OGIAB48TV
CG046479 PUFUV38TD

```
FEATURES
source
  Location/Qualifiers
  1..779
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBTa230M18"
  /clone_lib="ZM_0.6_1.0_KB"
  /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
  Cor selected genomic DNA library"

ORIGIN
Query Match      27.2%; Score 451; DB 28; Length 779;
Best Local Similarity 99.8%; Pred. No. 4.6e-203;
Matches 501; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 338 GCATCGGTGTTCTTTATTACATGAATAAATGACCGCTGCTTTTGCAATTAAGGGCTAGA 397
Db 502 GCATCGGTGTTCTTTATTACATGAATAAATGACCGCTGCTTTTGCAATTAAGGGCTAGA 443

QY 398 GATGTACGAGTGAACGCTTCGTAATTCATGACAGAGTAATTTTTTTTAATGACGGAG 457
Db 442 GATGTACGAGTGAACGCTTCGTAATTCATGACAGAGTAATTTTTTTTAATGACGGAG 383

QY 458 GCGCTTTGATTCCTCAGCAGCTTCGCGGCCCGTCTTCGTGTACGGCAGTGACGAGGC 517
Db 382 GCGCTTTGATTCCTCAGCAGCTTCGCGGCCCGTCTTCGTGTACGGCAGTGACGAGGC 323

QY 518 AGCAGCAGCTCTACGTGCCCCACGTGCTGCTGCCCTGCCACCGTGATGGCAAGGCTTC 577
Db 322 AGCAGCAGCTCTACGTGCCCCACGTGCTGCTGCCCTGCCACCGTGATGGCAAGGCTTC 263

QY 578 TTGCTTGCGTGCCTGCTGCTGGCAGCAGCAGAGTGAAGCTGACGGCGTGGCTGCTGCC 637
Db 262 TTGCTTGCGTGCCTGCTGCTGGCAGCAGCAGAGTGAAGCTGACGGCGTGGCTGCTGCC 203

QY 638 TCGCGCGTGTAGCAGCGCAGCGTGCAGCTGATGCTAGCGTTGCTACTTTTGTGT 697
Db 202 TCGCGCGTGTAGCAGCGCAGCGTGCAGCTGATGCTAGCGTTGCTACTTTTGTGT 143

QY 698 TTGTTTTCTCTGTGCTTGGATGATGACGTAAATAAAGCTGTGTATCTCATGTGATTGA 757
Db 142 TTGTTTTCTCTGTGCTTGGATGATGACGTAAATAAAGCTGTGTATCTCATGTGATTGA 83

QY 758 TCGACGTGTGCGCGAATTTCAATCAGTAAACAATCAGTGCATCCCATCGACTTCG 817
Db 82 TCGACGTGTGCGCGAATTTCAATCAGTAAACAATCAGTGCATCCCATCGACTTCG 23

QY 818 TGTGCGCGATAAATCTGCTAGG 839
Db 22 TGTGCGCGATAAATCTGCTAGG 1

RESULT 2
CG152000/c
LOCUS
DEFINITION
  728 bp DNA linear GSS 21-AUG-2003
  PUIFT17TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0573C10,
  genomic survey sequence.
ACCESSION
  CG152000
VERSION
  CG152000.1 GI:34042783
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
    1 (bases 1 to 728)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
  Bennetzen, J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUIFT17TB

CG152000
LOCUS
DEFINITION
  498 bp DNA linear GSS 21-AUG-2003
  PUIFT17TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0573C10,
  genomic survey sequence.
ACCESSION
  CG151999
VERSION
  CG151999.1 GI:34042783
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
    1 (bases 1 to 498)
REFERENCE
  Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
  Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
  Bennetzen, J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUIFT17TB
```

Contact: Cathy Whitelaw
TIGR
712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seg primer: TP
Class: sheared ends.
Location/Qualifiers
1..728
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0573C10"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN

Query Match 26.3%; Score 436; DB 29; Length 728;
Best Local Similarity 99.5%; Pred. No. 6.4e-196;
Matches 656; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 601 AGCGACGAGTTGAAGCTACGGCGTGGCTTGCTTGGCTCGCGCGCTGCTAGCAGCGGAC 660
Db 728 AGCGACGAGTTGAAGCTACGGCGTGGCTTGCTTGGCTCGCGCGCTGCTAGCAGCGGAC 669

QY 661 GACGTGCGAGCTGATGCTAGCGTTGCTACTTTTGTGTTTGTGTTTCTCTGCTTGGGATG 720
Db 668 GACGTGCGAGCTGATGCTAGCGTTGCTACTTTTGTGTTTGTGTTTCTCTGCTTGGGATG 609

QY 721 ATGCGATTAATAAAGCTGTGATCTCATGTGATTCATCGCTGTCGGCGATTTCAA 780
Db 608 ATGCGATTAATAAAGCTGTGATCTCATGTGATTCATCGCTGTCGGCGATTTCAA 549

QY 781 ATCAGTAAACAATCACATGTCATCCCATCGACTTCGCTGTCGGCGATAAATCTGCTAGG 840
Db 548 ATCAGTAAACAATCACATGTCATCCCATCGACTTCGCTGTCGGCGATAAATCTGCTAGG 489

QY 841 ATTGTGCTTTTCATGTATGCTGTGAATAGAGAGAAAGATAATATTACATAAATAAT 900
Db 488 ATTGTGCTTTTCATGTATGCTGTGAATAGAGAGAAAGATAATATTACATAAATAAT 429

QY 901 GAGTTCAACATCTTTCTTAATACCTTTTCTCGATCTCCTTAGTAGAGCACCTGCTGTA 960
Db 428 GAGTTCAACATCTTTCTTAATACCTTTTCTCGATCTCCTTAGTAGAGCACCTGCTGTA 369

QY 961 TAACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCCAA-GGCTGCTTAAGTAAATA 1019
Db 368 TAACGTGTTAAGAACCCCTTGTACCGAGTGTAGTCCAAAGGGCTGCTTAAGTAAATA 309

QY 1020 GTAGAACACAGAGACAAGAGTAGAGAGAACTGATCTTTGTACTATATGCTGCT 1079
Db 308 GTAGAACACAGAGACAAGAGTAGAGAGAACTGATCTTTGTACTATATGCTGCT 249

QY 1080 GCTCTCCAAAGTTACATGATATGGGATCTCCTCTCTATTTATAGACAAAACCTAGGTT 1139
Db 248 GCTCTCCAAAGTTACATGATATGGGATCTCCTCTCTATTTATAGACAAAACCTAGGTT 189

QY 1140 TCAGGCATATGGGCCACATAGGCTTTCTGGCCCAAGAAAGGTTTCTTAACACTACATC 1199
Db 188 TCAGGCATATGGGCCACATAGGCTTTCTGGCCCAAGAAAGGTTTCTTAACACTACATC 129

QY 1200 TTCAGGCGGCTGCTGCTGCTTCAACTCTTCCCTCCGCTCCAGAGCGGATAAATA 1258
Db 128 TTCAGGCGGCTGCTGCTTCAACTCTTCCCTCCGCTCCAGAGCGGATAAATA 70

RESULT 3

CG151999
LOCUS
DEFINITION

```

genomic survey sequence.
CG151999
CG151999.1 GI:34042782
GSS.
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 498)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PU1F7L7TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..498
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0573C10"
/clone_lib="ZM 0.6 1.0 kb"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN
Query Match 22.6%; Score 374; DB 29; Length 498;
Best Local Similarity 99.8%; Pred. No. 2.2e-166;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 338 GCATGCGTGTCTTTATTACATGAAAAAATGAGCGCTGCTTTTGCAATTAAGGGTCTAGA 397
Db 74 GCATGCGTGTCTTTATTACATGAAAAAATGAGCGCTGCTTTTGCAATTAAGGGCTAGA 133

QY 398 GATGTACAGTGCAACGCTTCGTATTCACCTCACAGTAATTTTATATGACGGAG 457
Db 134 GATGTACAGTGCAACGCTTCGTATTCACCTCACAGTAATTTTATATGACGGAG 193

QY 458 GCGCTTTGTATTCCTCAGCAGCTCTGCGCGCCGCTTCTGTAGCGCAGTGACGAGGC 517
Db 194 GCGCTTTGTATTCCTCAGCAGCTCTGCGCGCCGCTTCTGTAGCGCAGTGACGAGGC 253

QY 518 AGCAGCAGCTTACGTGCCACGCTGCTGCTGCTGCCACCGTATGGAAGGCTTGC 577
Db 254 AGCAGCAGCTTACGTGCCACGCTGCTGCTGCCACCGTATGGAAGGCTTGC 313

QY 578 TTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
Db 314 TTGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373

QY 638 TGGCGCGTGTAGCAGCGGACGACGCTGACGCTGATGCTAGCGTGTGCTACTTTTGT 697
Db 374 TGGCGCGTGTAGCAGCGGACGACGCTGACGCTGATGCTAGCGTGTGCTACTTTTGT 433

QY 698 TTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
Db 434 TTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493

QY 758 TCGAC 762
Db 494 TCGAC 498

FEATURES
source
1..754
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0639C04"
/clone_lib="ZM 0.6 1.0 kb"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN
Query Match 18.9%; Score 313; DB 29; Length 754;
Best Local Similarity 99.7%; Pred. No. 1.9e-137;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1252 ATAAATACGGGCTCTACGGTGCACGCGCATGTTGCGCGGAGGGCATCTCTGCCCG 1311
Db 364 ATAAATACGGGCTCTACGGTGCACGCGCATGTTGCGCGGAGGGCATCTCTGCCCG 305

QY 1312 GGGATGCGCGCAGCGCATGCTCGACACGACGATGTTGCGCGGAGGGCTCTCT 1371
Db 304 GGGATGCGCGCAGCGCATGCTCGACACGACGATGTTGCGCGGAGGGCTCTCT 245

QY 1372 CGCACAGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1431
Db 244 CGCACAGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185

QY 1432 CAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1491
Db 184 CAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 125

QY 1492 CCCACCGAGCGGTATATAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1551
Db 124 CCCACCGAGCGGTATATAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65

QY 1552 AGCAGCAACAGCCACCACTAACTAAGAACAGTAGTAGTCCCTGTGTAGCTTAGCAA 1611
Db 64 AGCAGCAACAGCCACCACTAACTAAGAACAGTAGTAGTCCCTGTGTAGCTTAGCAA 5

QY 1612 CGAT 1615
Db 4 CGAT 1

```

RESULT 4

CG146202/c

LOCUS

DEFINITION

CG146202

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 754)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other_GSSs: PUJCVI47D

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..754

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTa0639C04"

/clone_lib="ZM 0.6 1.0 kb"

/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

CoT selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity

Matches 363; Conservative

0; Mismatches 1; Indels

0; Gaps

0;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db


```

ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS      1 (bases 1 to 949)
              Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL
COMMENT      Unpublished (2003)
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TP
              Class: sheared ends.
FEATURES
source       1..949
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone="ZMMBTA127G13"
              /clone_lib="ZM 0.6-1.0 KB"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"
ORIGIN
Query Match      15.0%; Score 248; DB 28; Length 949;
Best Local Similarity 100.0%; Pred. No. 1.4e-106;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTCACGGCTCAATACACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT 60
Db 409 GAATTCACGGCTCAATACACAGTCATCTACATGTGATAAATTTTATAGTTAGATTACTT 350
QY 61 CTTCGTATCATTTTCAGAGGTGAAAAAATCCAGAAAGCAAAATATTTTAAATGAAT 120
Db 349 CTTCGTATCATTTTCAGAGGTGAAAAAATCCAGAAAGCAAAATATTTTAAATGAAT 290
QY 121 GATGCAATATACAAATTTAATTACAAATATATGTAAGATTACATGTTTATGTTTCATAGA 180
Db 289 GATGCAATATACAAATTTAATTACAAATATATGTAAGATTACATGTTTATGTTTCATAGA 230
QY 181 AATCAATTTCTAGAGTCATAATATGCTAAATTTAAATTTCTATTTCTTTAAAG 240
Db 229 AATCAATTTCTAGAGTCATAATATGCTAAATTTAAATTTCTATTTCTTTAAAG 170
QY 241 AAAAAAAA 248
Db 169 AAAAAAAA 162
RESULT 8
CG164746/c
LOCUS
DEFINITION     CG164746 728 bp DNA linear GSS 21-AUG-2003
                PUII051TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTA0593105,
                genomic survey sequence.
ACCESSION
VERSION        CG164746.1 GI:34055547
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS      1 (bases 1 to 728)
              Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium

```

```

JOURNAL
COMMENT      Unpublished (2003)
              Other GSSs: PUII051TD
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TP
              Class: sheared ends.
FEATURES
source       1..728
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone="ZMMBTA0593105"
              /clone_lib="ZM 0.6-1.0 KB"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"
ORIGIN
Query Match      5.2%; Score 86; DB 29; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.7e-29;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1000 GGCTGCTTAAGTGAATATAGTAGACACACAGGAGACAGAGTGTAGAGAGGAACTGATT 1059
Db 669 GGCTGCTTAAGTGAATATAGTAGACACACAGGAGACAGAGTGTAGAGAGGAACTGATT 610
QY 1060 CTTTGTACTATATGCTGCTCTCTC 1085
Db 609 CTTTGTACTATATGCTGCTCTCTC 584
RESULT 9
CG062076/c
LOCUS
DEFINITION     CG062076 786 bp DNA linear GSS 19-AUG-2003
                PUPPD96TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBTA0701024,
                genomic survey sequence.
ACCESSION
VERSION        CG062076.1 GI:33934256
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS      1 (bases 1 to 786)
              Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
TITLE        Maize Genomics Consortium
JOURNAL
COMMENT      Unpublished (2003)
              Other GSSs: PUPPD96TB
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TP
              Class: sheared ends.
FEATURES
source       1..786
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone="ZMMBTA0701024"
              /clone_lib="ZM 0.6-1.0 KB"
              /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"

```

```

ORIGIN
Query Match          4.0%; Score 67; DB 29; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.8e-20;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAGAGTGATAGAGGAAGTCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 1096
    |||||||
Db 519 AAGAGTGATAGAGGAAGTCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 460
    |||||||
QY 1097 TGATATG 1103
    |||||||
Db 459 TGATATG 453

RESULT 10
CG062075
LOCUS
DEFINITION
  PURPD96TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTA0701024,
  genomic survey sequence.
ACCESSION
CG062075
VERSION
CG062075.1 GI:33934255
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 858)
AUTHORS
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other GSSs: PUPPD96TD
  Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
  1..858
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBTA0701024"
  /clone_lib="ZM_0.6_1.0_KB"
  /notes="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
  Cot selected genomic DNA library"

FEATURES
  source
  1..858
  |||
  22 GCATG 19

ORIGIN
Query Match          4.0%; Score 67; DB 29; Length 858;
Best Local Similarity 100.0%; Pred. No. 1.7e-20;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAGAGTGATAGAGGAAGTCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 1096
    |||||||
Db 444 AAGAGTGATAGAGGAAGTCTTTGTTACTATATGCTGCTCTCCAAAGGTTACA 503
    |||||||
QY 1097 TGATATG 1103
    |||||||
Db 504 TGATATG 510

RESULT 11
CG193723/c
LOCUS
DEFINITION
  PURPD96TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTA0642H02,
  genomic survey sequence.

```

```

ACCESSION
CG193723
VERSION
CG193723.1 GI:34084784
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 438)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other GSSs: PUJDJ37TD
  Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
  1..438
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBTA0642H02"
  /clone_lib="ZM_0.6_1.0_KB"
  /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
  Cot selected genomic DNA library"

FEATURES
  source
  1..438
  |||
  22 GCATG 19

ORIGIN
Query Match          3.9%; Score 64; DB 29; Length 438;
Best Local Similarity 100.0%; Pred. No. 6e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1083 CTCCTAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 1142
    |||||||
Db 82 CTCCTAAGGTTACATGATGGGATCTCTCTCTATTATAGACAAACTAGGGTTTCA 23
    |||||||
QY 1143 GCATG 1146
    |||
  22 GCATG 19

RESULT 12
CG193725
LOCUS
DEFINITION
  PUJDJ37TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTA0642H02,
  genomic survey sequence.
ACCESSION
CG193725
VERSION
CG193725.1 GI:34084786
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 446)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other GSSs: PUJDJ37TB
  Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

```

Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
Source
1..446
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA0642H02"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 29; Length 446;
Best Local Similarity 100.0%; Pred. No. 5.9e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1083 CTCCAAAGGTTACATGATGGGATCTCCTCTCTATTATTATAGACAAACTAGGTTTCA 1142
Db 357 CTCCAAAGGTTACATGATGGGATCTCCTCTCTATTATTATAGACAAACTAGGTTTCA 416

QY 1143 GGCA 1146
Db 417 GGCA 420

RESULT 13
CC352433/c
LOCUS
DEFINITION
FUHDX59TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA429I22,
genomic survey sequence.
ACCESSION
CC352433
VERSION
CC352433.1 GI:30821833
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 732)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
Source
1..732
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA429I22"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 28; Length 732;
Best Local Similarity 100.0%; Pred. No. 4.9e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1083 CTCCAAAGGTTACATGATGGGATCTCCTCTCTATTATTATAGACAAACTAGGTTTCA 1142
Db 357 CTCCAAAGGTTACATGATGGGATCTCCTCTCTATTATTATAGACAAACTAGGTTTCA 416

QY 1143 GGCA 1146
Db 417 GGCA 420

RESULT 14
CC006284
LOCUS
DEFINITION
PUD145TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA219G20,
genomic survey sequence.
ACCESSION
CC006284
VERSION
CC006284.1 GI:29384846
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 884)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
Source
1..884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA219G20"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 3.9%; Score 64; DB 28; Length 884;
Best Local Similarity 100.0%; Pred. No. 4.6e-19;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1130 AACTAGGGTTTCAGGCATATGGCCACATAGGCCCTCTCTGCGCCCAAGAGGTTTCTTAA 1189
Db 812 AACTAGGGTTTCAGGCATATGGCCACATAGGCCCTCTCTGCGCCCAAGAGGTTTCTTAA 871

QY 1190 CACT 1193
Db 872 CACT 875

RESULT 15
CG096594
LOCUS
DEFINITION
PUFUD51TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0733J10,
genomic survey sequence.
ACCESSION
CG096594
VERSION
CG096594.1 GI:33978888
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 844)

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
Bennetzen, J.

TITLE

Maize Genomics Consortium
Unpublished (2003)

COMMENT

Other GSSs: PUFUD53TD
Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

FEATURES

source

Location/Qualifiers

1..844

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZM0610733J10"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high"

Cot selected genomic DNA library"

ORIGIN

Query Match

Best Local Similarity 3.6%; Score 60; DB 29; Length 844;

Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1033 AGACAGAGCTGTAGAGAGAACTGATTCCTTGTACTATATGCTGCTCTCCAAAGGT 1092

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 476 AGACAGAGCTGTAGAGAGAACTGATTCCTTGTACTATATGCTGCTCTCCAAAGGT 535

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1093 TACATGATATGGGATCTCTCTCTATTTATAGACAAAAGTGGGTTTCAG 1143

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 536 TACATGATATGGGATCTCTCTCTATTTATAGACAAAAGTGGGTTTCAG 586

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: September 2, 2004, 14:41:27

Job time : 3930 secs